BINDING SERVICES
Tel +44 (0)29 2087 4949
Fax +44 (0)29 20371921
e-mail bindery@cardiff.ac.uk

Cross-Sectional Atomic Force Microscopy of III-V Semiconductor Device Structures

Christian Jenkins, MPhys (Wales) MPhil (Wales)

A thesis submitted for the degree of Doctor of Philosophy
in the Department of Physics and Astronomy,
University of Wales, Cardiff.

January 2004

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Dedicated to the memory of my late Grandfather, Elidor Lewis

Contents

Chapter 3

The Oxidation of III-V Semiconductor Materials

Chapter 4

Atomic Force Microscopy

Chapter 7

Studies of Device Structures

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed .....(candidate)
Date 24/6/04

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed .....(candidate)
Date 24/6/04

Statement 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .....(candidate)
Date 24/6/04

Acknowledgements

Firstly, I must thank my supervisor, Dr. David Westwood for his support and encouragement during my studies.

Much gratitude must also be extended to Martin Elliott, Emyr Macdonald, Paul Hulyer, Huw Summers, Gigi Iordache, Angela Kestle, Ian Robinson and Rob Tucker; all of who provided some form of expertise along the way. Thanks also go to the staff of the physics workshop, Glyn Summers, Dave Pickering and Jeff Trivett for their technical prowess and sunny dispositions!

This research would not have been possible without the collaboration of Clive Meaton, Su Sajip and Steve Bland at IQE, and I thank them for the opportunity of carrying out the work.

For making it such an enjoyable time in Cardiff, I would like to thank everyone in the Physics Department, particularly John Thompson, Louise Haworth, and Nick & Alison Tripp. A special mention must, of course go to those at 11 Cosmeston Street for providing ample amusement and friendship: Paul Jones (and those late night trips to Tesco), Emma Phillips (and her bizarre home-made wine) and Loretta Dunne (and her hang-glider).

A special mention must also go to those friends outside the world of physics, particularly Laura Palmer, Philip Harris, Joanne Deane, Craig Roberts, Martha Davison, Meryl Thomas and Brian Dowric, for providing copious amounts of distraction. Also the field-trippers of the Glamorgan Bird Club, particularly Geri Thomas and Neil Donaghy, for those relaxing trips to wide, open (and often wet and miserable) spaces. Special thanks must also be expressed to Hayley Thomas, for providing me with the final motivation for getting it all done.

Finally, love and gratitude must go to my parents, who have always given me so much support and encouragement - I couldn't have done it without them.

I wish to acknowledge the financial support of Engineering and Physical Sciences Research Council (EPSRC) and International Quantum Epitaxy (IQE), and Digital Instruments for kind permission to print some of their diagrams.

Abstract

Atomic force microscopy (AFM) in air has been used to study various III-V semiconductor heterostructures. Topography of the (110) cleaved cross-sections has been examined where oxidation processes modify the surface and allow the structures to be investigated. This research aims to establish the potential of this technique as a metrology tool for use in an industrial environment. $Al_xGa_{1-x}As/GaAs$ was used as the prototypical system, and a test structure grown in order to establish how differences in oxidation rates between the different material compositions may be used to establish composition and layer thickness for heterostructure devices. The dependence of oxidation rate on layer composition and thickness has been confirmed. The mechanisms of field-aided and diffusion limited growth have been determined to be responsible for the oxidation of the $Al_xGa_{1-x}As$ layers within the test structure, with field aided being the dominant mode for materials $x < 0.8$ and diffusion limited dominating for layers $x > 0.8$ observed. In addition, a new effect of interface enhanced oxide growth have been observed and quantified in terms of layer composition. It is found to be most important for $Al_{0.9}Ga_{0.1}As$ layers of ~ 50 - 100 nm thickness during relatively early stages of oxidation. The use of phase detection microscopy has also been applied for the first time to determine the presence of layers where no measurable step heights are present. Unlike previous reports, there has been no observed difference in oxidation rate between *p*- and *n*-type materials. These findings have been applied to real device structures, where the material composition of $Al_xGa_{1-x}As$ can be determined to within $x \pm 0.02$ and layer widths may be determined to within 3 nm at best. It has been shown that step height differences of as little as 0.1 Å are sufficient to distinguish between layers, and that quantum wells of as little as 42 Å in width are detectable.

Chapter 1

General Introduction

1.1 Introduction

Current commercial interest in semiconductor materials is very high indeed. The demands of the modern world require constant advances in device performance. With applications ranging from telecommunications to computer systems and solar cells, there is little doubt that the microelectronics revolution has changed the world forever. In particular, optical semiconductor devices, such as the light-emitting diode and semiconductor laser are becoming increasingly important with the growing trend to use optical fibres for communication.

At the heart of device research lies the characterisation techniques. Most of them are well established, such as optical microscopy, X-ray diffraction and scanning electron microscopy (SEM). However, within the last twenty years or so, exciting scanning probe microscopies, such as scanning tunnelling microscopy and atomic force microscopy have emerged [1], that have greatly expanded the knowledge of solid surfaces in general. The power now exists to probe properties on the atomic scale [2].

Compared with other techniques, scanning probe microscopy is still in its infancy, but its use in industry as an aid to device characterisation is becoming more widespread. This has been particularly helped by companies such as Digital Instruments who, with the development of their series of ever improving NanoScope^{TM}s, have introduced 'user friendly' scanning tunnelling and atomic force capabilities.

1.2 Areas of Study

The precise measurement of all the individual layer thicknesses in complex epitaxially grown device structures has always been difficult [3]. Optical or X-ray

diffraction techniques only give reliable values for the simplest structures, and more direct methods such as transmission electron microscopy or scanning electron microscopy can be expensive and very time consuming. However, with the advent of atomic force microscopy, which can operate in air whilst maintaining near atomic resolution, there now exists the possibility of quick, simple, accurate and precise thickness measurements on even the most complex structures.

Here studies were carried out using a Digital Instruments NanoScope™ IIIa MultiMode™ SPM operating under TappingMode™ SPM. The approach involved the cleaving of an epitaxial device structure to reveal its (110) cross-section in air followed by the measurement of its topography to measure the layer dimensions. Contrast between the layers resulted from the different oxidation rates for different material compositions [4, 5].

Important questions addressed were: how accurately layer thicknesses and compositions could be determined, the minimum oxide step height measurable and what time-scale was involved in obtaining data. The effects of oxide thickness, varied by the exposure time to air and compositional dependent oxidation rates, on surface morphology and consequently the ability to determine the true position of layer interfaces was considered.

The primary goal of this research was to assess the suitability of this technique as a characterisation tool for optoelectronic semiconductor devices. AlGaAs/GaAs structures were chosen as the prototypical system. Here, the relatively high oxidation rate of AlGaAs compared with other III-V materials allowed for an effective way of refining the technique. An AlGaAs/GaAs test structure was grown by metal organic vapour phase epitaxy (MOVPE) by International Quantum Epitaxy Ltd. (IQE), St. Mellons, Cardiff. This sample was extensively studied in order to establish oxidation rates and optimum scan conditions for the material. An added advantage is that AlGaAs based devices are in widespread use, so comparison with real devices could be made. Device structures investigated included vertical cavity surface emitting lasers (VCSELs), laser diodes and light emitting diodes (LEDs). In addition, a phosphide based multiple quantum barrier test structure was also investigated to determine the technique's applicability to phosphide based materials.

1.3 Layout of Thesis

Chapter 2 contains a resume of the physics concerning the design and operation of semiconductor optoelectronic devices.

Chapter 3 is a discussion of the oxidation of III-V materials outlining the present understanding taken from the literature. Clearly, an understanding of the oxidation processes is of vital importance to this research. Unfortunately, as is demonstrated in this chapter, no clear theory for their detailed oxidation process exists.

Chapter 4 details atomic force microscopy, and in particular refers to the Digital Instruments NanoScope™ IIIa MultiMode™ SPM, which was used exclusively throughout this study.

The results of this study are presented in chapters 5, 6 and 7. Chapter 5 outlines the experimental considerations, such as sample preparation and sources of error. Chapter 6 deals with the detailed measurements that were carried out on the test structure. The validity of these results were then investigated through real device structures, the results of which are presented in chapter 7.

The thesis concludes with a discussion and suggestions for future work in chapter 8.

1.4 References

[1] *MultiMode™ SPM Instruction Manual.* Digital Instruments.

[2] R J Colton, A Engel, J E Frommer, H E Gaub, A A Gewirth, R Guckenberger, J Rabe, W M Heckl & B Parkinson (Eds.). *Procedures in Scanning Probe Microscopies.* John Wiley and Sons (1998).

[3] A Rudra, L Sagalowicz, K Leifer, J Behrend, C A Berseth, O Dehaese & J F Carlin. *Journal of Crystal Growth.* **188**, 300 (1998).

[4] F Reinhardt, B Dwir, G Biasiol & E Kappon. *Applied Surface Science,* **104/105**, 529 (1996).

[5] C Jenkins, D I Westwood, M Elliott, J E Macdonald, C Meaton & S Bland. *Materials Science and Engineering B.* **80**, 138 (2001).

Chapter 2

III-V Semiconductor Optoelectronic Devices

2.1 Introduction

There are a number of reasons why silicon is still the world's most widely used semiconductor material. It is relatively easy to fabricate in pure form, has excellent processing properties and has a high quality native oxide. III-V materials on the other hand generally have poor quality oxides, and are more expensive to manufacture. The interest in III-V materials lies in the fact that they exhibit far superior optical properties than silicon, and that electron mobility is higher, allowing the production of optoelectronic and high frequency devices.

Following on from the theoretical work of W P Dumke, scientists at IBM, GE and Bell Labs demonstrated that III-V materials could lase in the early 1960s [1, 2, 3, 4]. However, the above mentioned processing challenges have meant that the III-V market has not taken off until relatively recently. Indeed, improvements in fabrication have now lead to III-Vs being utilised in other semiconductor devices, such as the heterojunction bipolar transistor (HBT), an integral component of the mobile phone. Optoelectronic devices, such as the laser diode and the vertical cavity surface emitting laser (VCSEL) are described later in this chapter.

While the primary concern of this research is not with the operation of the devices, an understanding of their design and operation is desirable. Therefore, this chapter gives a brief description of the design and operation of III-V semiconductor devices. It is of an intentionally qualitative nature, the aim being to merely endow the reader with a better understanding of the complexity of their design and an appreciation for the motivation behind this work. For a more thorough treatment of the material, the reader is directed toward one of a number of books specialising in semiconductor devices (for example, Sze [5] or Singh [6]).

Following a basic introduction to the properties of semiconductors, the specific reasons why III-Vs are used extensively in the optoelectronics industry are considered. A discussion of the various growth techniques and issues associated with

device substrates follows. The various optoelectronic devices that have been encountered in this work will then be considered. Beginning with the *pn* junction, these progress in complexity to the VCSEL.

2.2 Optical Properties of Semiconductor Materials

At any temperature greater than absolute zero, there is a probability that electrons may be thermally excited from the valence band and into the conduction band of a semiconductor material. However, an electron may also be excited across the bandgap optically, as illustrated in fig 2.1. Here, an electron absorbs the energy of a photon, whereupon it is promoted to the conduction band, leaving behind a hole in the valence band. The process will only occur if the photon provides the electron with sufficient energy to overcome the bandgap energy of the semiconductor. Since the energy of a photon with frequency ν and wavelength λ is given by

$$E = h\nu = \frac{hc}{\lambda}, \tag{2.1}$$

the minimum frequency, ν_{min}, and corresponding maximum wavelength, λ_{max}, at which absorption can occur is given by

$$\nu_{min} = \frac{c}{\lambda_{max}} = \frac{E_g}{h} \tag{2.2}$$

Therefore, exposing a semiconductor to light of the appropriate frequency may significantly increase the conductivity of the material. This process is known as *photoconductivity*.

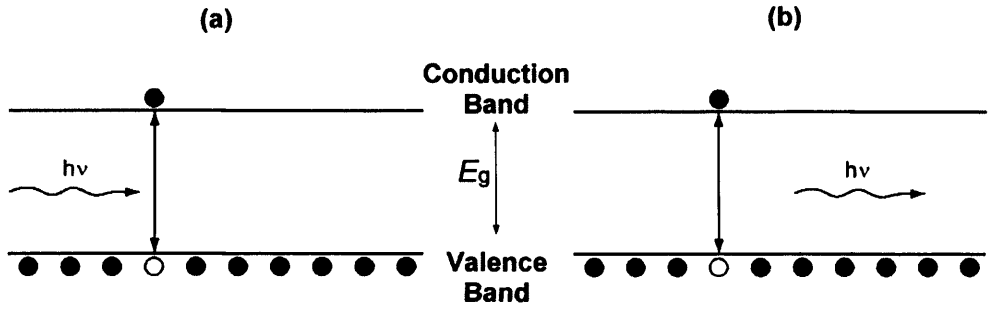

Figure 2.1 *(a) A single photon is capable of exciting an electron from the valence band into the conduction band provided that the energy of the photon is greater than the bandgap energy E_g of the semiconductor.* ***(b)*** *Conversely, a conduction electron can recombine with a hole and produce a photon [7].*

The reverse process may also occur whereby an electron in the conduction band may recombine with a hole in the valence band, with the energy lost being emitted as a photon.

This process is the basis for the operation of optoelectronic devices, and although this treatment has simplified greatly what is in fact a very complicated process, the knowledge is sufficient to understand their operation.

There is one complication, however, that must be mentioned here. Not all semiconductor materials are capable of producing light efficiently. This is due to the fact that the properties of electrons and holes within the crystal are specified by the energy E and the wavevector $\boldsymbol{k}$ (or '*crystal momentum*'). Recombination of an electron and hole can only occur if both energy and momentum are conserved in the process.

Energy may be conserved by a photon of the correct frequency. However, since a photon has negligible momentum, it will have little affect on the momentum of the electron. In materials where the conduction band minimum and valence band maximum coincide in momentum, transitions may occur with the aid of a photon. These so-called *direct gap* materials have excellent optical properties since the likelihood of recombination is high. Semiconductors whose conduction band minimum and valence band maximum do not line up in momentum are referred to as *indirect gap* materials. Optical transitions are possible in these materials, but require a phonon in order for momentum to be conserved. Consequently, recombination is

much less likely in these materials, resulting in poor optical properties. These recombination processes are illustrated in fig 2.2.

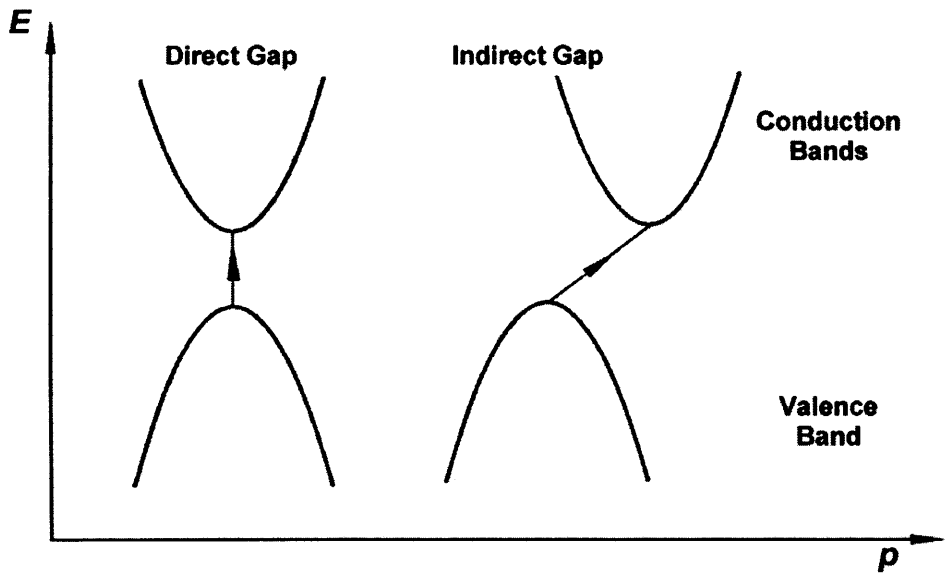

Figure 2.2 An energy/momentum plot illustrating the difference between direct and indirect gap transitions in a semiconductor.

2.3 III-V Semiconductor Material Properties

The reason silicon has such poor optical properties is that it is an indirect bandgap material (fig 2.3). On the other hand, many of the III-V materials have direct bandgap regimes, and for this reason they are now extensively used in the production of optoelectronic devices. Of the III-Vs, gallium arsenide (GaAs) is perhaps the most important and widely used semiconductor. The direct bandstructure of GaAs is illustrated in fig 2.3. Coupled with the excellent optical properties, is the fact that GaAs has a relatively low effective electron mass that leads to superior electron transport in the conduction band [5]. The larger bandgap also ensures better high-temperature performance [6].

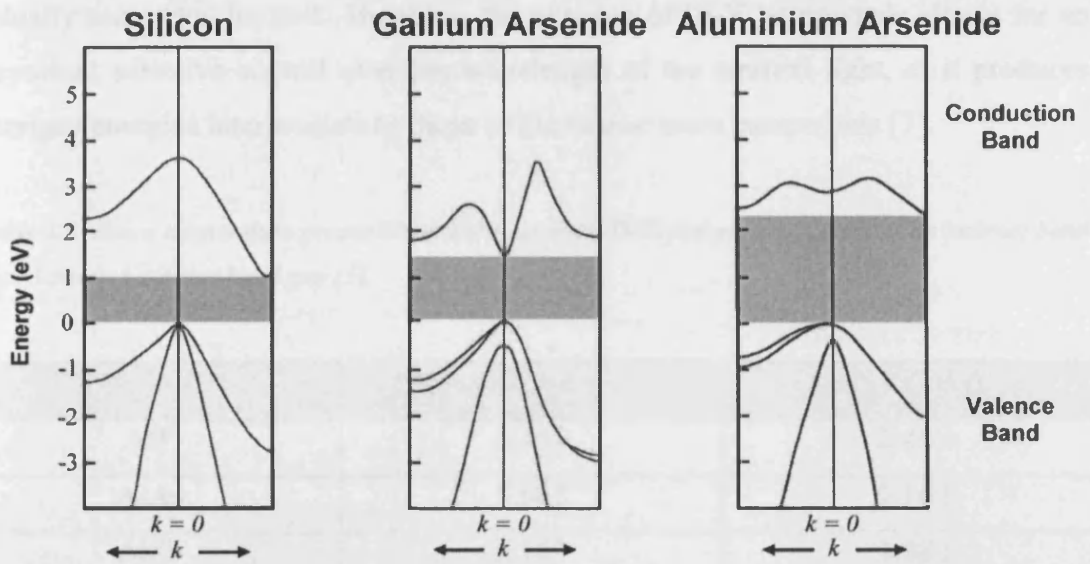

***Figure 2.3** Bandstructures of Silicon, Gallium Arsenide and Aluminium Arsenide [6].*

Another important III-V material is aluminium arsenide (AlAs). Despite having an indirect energy bandgap (see fig 2.3), AlAs has found an important use in alloy with GaAs (see section 2.4). Excellent lattice constant matching between the two has lead to the widespread use of GaAs/AlGaAs heterostructures [5].

Of the remaining III-V materials, two more that are commercially important should be mentioned. The first, indium arsenide (InAs), has a narrow bandgap with a very low electron mass, giving superior transport properties [6]. In order to achieve suitable lattice-matching, it is often used in alloy (e.g. InGaAs, InAlAs, etc.) for devices. Finally, indium phosphide (InP) is used as a barrier layer material for semiconductor lasers and as a substrate for devices compatible with low loss optical windows in optical fibres [6]. Its' alloys, such as InGaP and AlGaInP can be made to lattice match to GaAs substrates.

2.4 III-V Semiconductor Material Alloys

For III-V materials to be of any practical use, they must be available in a wide range of bandgap energies in order to produce lasers of the desired wavelength. A list of the properties of the most widely used III-V semiconductors appears in table 2.1. By referring to this data, it can be seen that the bandgap energies of the materials are

actually somewhat limited. However, the alloying of III-V compounds allows for an important selective control over the wavelength of the emitted light, as it produces bandgap energies intermediate to those of the two or more compounds [7].

Table 2.1 *Room temperature properties of some common III-V compounds. i denotes an indirect band gap, d denotes a direct band gap [5].*

Material	Lattice Parameter (Å)	Energy Gap (eV)
AlP	5.451	2.45 *i*
AlAs	5.661	2.16 *i*
AlSb	6.163	1.58 *i*
GaP	5.451	2.26 *i*
GaAs	5.653	1.42 *d*
GaSb	6.096	0.73 *d*
InP	5.869	1.35 *d*
InAs	6.058	0.36 *d*
InSb	6.479	0.17 *d*

As discussed later in this chapter, most III-V devices consist of a number of layers of different III-V materials. One of the restrictions in device design, is that the materials either side of a layer interface must have the same, or at least a very similar lattice constant (in this case, the materials are said to be *lattice matched*). If a device were produced where the lattice constants of the materials were very different, strain would be built up in the layers that could only relieve itself by the formation of large numbers of dislocations within the crystal lattice. In this case, the end result would be a defective device. It must be noted however, that in situations where the two materials are not lattice matched, it is still possible to grow good quality layers, provided the grown layer is fairly thin. Such thin layers have proved to be of value to the device designer, particularly in the formation of quantum wells [8].

Again, referring to table 2.1, it can be seen that the choice of lattice constants is also limited. Thankfully, the same alloying process that allows the ability to produce a desired bandgap, gives control over lattice constants by producing crystals with an atomic separation intermediate to that of the two materials in the alloy.

From this, it is obvious that the need for producing materials not only with the correct bandgap energy, but also of a suitable lattice constant complicates device design. Indeed, this complication has forced the development not only of so-called ternary compounds (i.e. those produced by alloying two different III-V materials), but of quaternary and even quinternary III-V compounds.

Of great use is the data plotted in fig 2.4. From it, the determination of what materials give the desired bandgap energy and lattice constant is possible. The lines joining each III-V compound give the values of lattice constant and bandgap energy that may be achieved by alloying those two materials, since the intermediate properties of the alloy lie along these lines. The properties of quaternary materials are difficult to determine from this diagram, but they lie within the area bounded by the lines joining the three III-V materials in question.

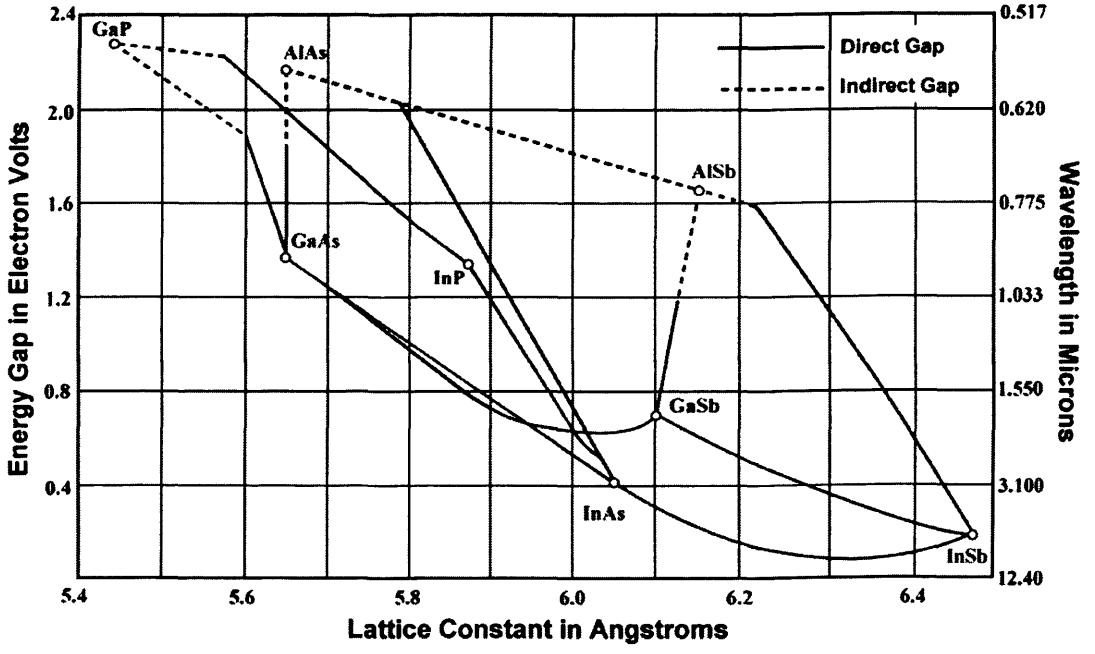

Figure 2.4 Energy bandgap and lattice constants for various III-V semiconductors. Lines joining the binary materials indicate ternary materials with their corresponding emission wavelengths as shown on the right. Dashed lines indicate indirect bandgap materials, which are unsuitable for laser operation [6].

It is immediately apparent from fig 2.4 that not all of the lines joining the III-V compounds are straight, indicating the variation in the properties of the alloys is not

necessarily a linear one. In the case of the alloys of AlAs and GaAs, the transition from GaAs to AlGaAs through to AlAs lies along a near vertical line. This is due to the fact that the lattice constants of both AlAs and GaAs are almost identical. If the lattice constants of the two materials are significantly different however, an increase in disorder within the crystal creates an enhanced non-linear variation in the properties of the alloy. This manifests itself in an increased 'bowing' of the line joining the two materials, as is observed with most of the materials in fig 2.4 [5].

Since the proportion of materials within an alloy have such a great influence over its properties, it is important to be able to quantify the proportion of each constituent. Ternary compounds, such as AlGaAs are written in the form $Al_xGa_{1-x}As$. The subscript x indicates the mole fraction of AlAs within the alloy, while $(1-x)$ indicates the amount of GaAs. The value of x ranges between 0 (pure GaAs) and 1 (pure AlAs), although mole fractions are sometimes quoted as a percentage. If we consider that As is common to both compounds, it can be seen that x is also a measure of the Al mole fraction with respect to Ga.

2.5 Device Growth Techniques

Growth techniques have played an integral role in the rapid development of III-V semiconductor technology. Indeed, if it were not for such techniques, it would not be at the stage of advancement it is today. Epitaxial crystal growth is now so refined that it is possible to control such parameters as material composition to monolayer dimensions. This allows the creation of the very abrupt interfaces that contribute to the efficiency of devices. Constant refinement over the years has also allowed the reduction of the dimensions of the devices produced, and very sophisticated semiconductor heterostructures are now possible.

Since the early 1980s, two growth techniques have emerged as the leading technologies for the mass-production of advanced III-V devices [9]: *metal organic vapour phase epitaxy* (MOVPE) and *molecular beam epitaxy* (MBE). Both MBE and MOVPE are discussed below. Other growth techniques exist, but most are essentially variants of the above, and shall not be considered here.

2.5.1 Molecular Beam Epitaxy (MBE)

A schematic diagram of a typical molecular beam epitaxy system is shown in fig 2.5. The basis of the technique is to allow beams of desired atoms or molecules (e.g. As_4) to fall upon a substrate, where they react to form an epitaxial film. The substrate is placed in an ultrahigh-vacuum (UHV) (typically 10^{-10} Torr [9]) held at an elevated temperature (typically 400°C [9]). The temperature is chosen as to allow optimum mobility of atoms on the surface, ensuring they can locate good growth sites.

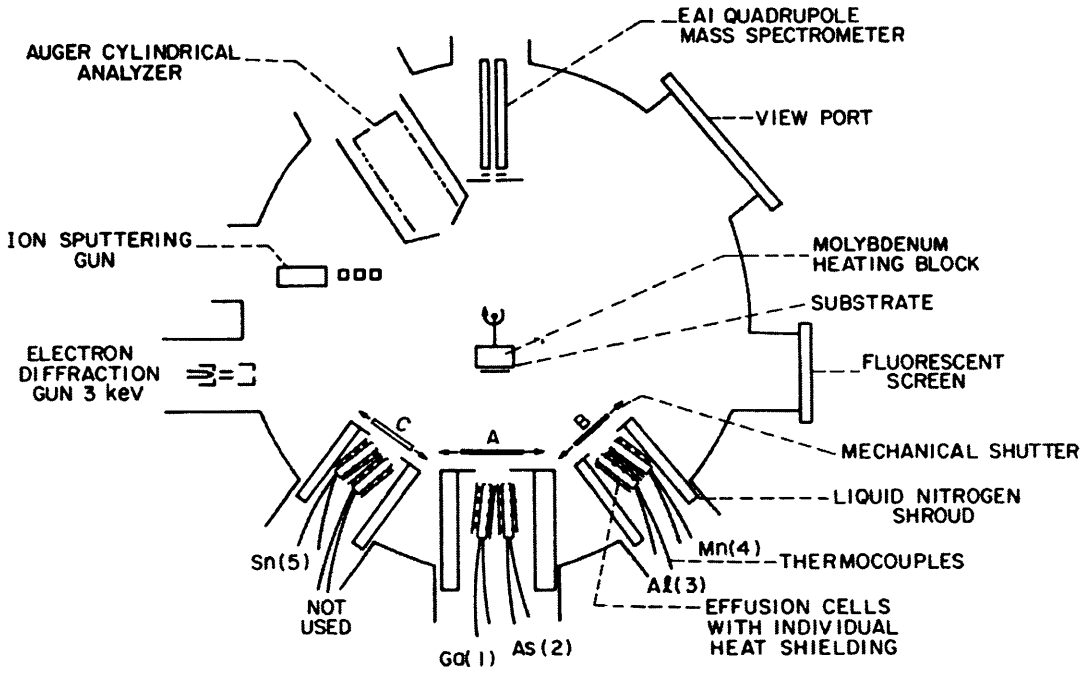

Figure 2.5 A schematic illustration of an MBE system [9].

The sources of the molecular beams are called *effusion cells*. They are essentially heated crucibles that contain the elemental sources required for the molecular beam. In the case where layers of a compound semiconductor are to be grown, separate cells are required for each of the component elements (e.g. Ga, In, Al or As etc.). Growth rates of 0.01 - 0.1 nm s^{-1} [10] are readily attained, while molecular beam fluxes are controlled by the temperature of the effusion cells. In a similar fashion, appropriate dopants may also be introduced into layers.

2.5.2 Metal Organic Vapour Phase Epitaxy (MOVPE)

A schematic diagram of the essential features of a typical metal organic vapour phase epitaxy system is shown in fig 2.6. MOVPE has proved itself to be a versatile and reliable technique for growing a wide variety of III-V semiconductor materials. As with MBE, the materials produced display excellent optical and electrical quality. Precise compositional control is readily achieved with both ternary and quaternary materials and the position of dopant layers can be controlled to within about 2 or 3 atomic layers [11].

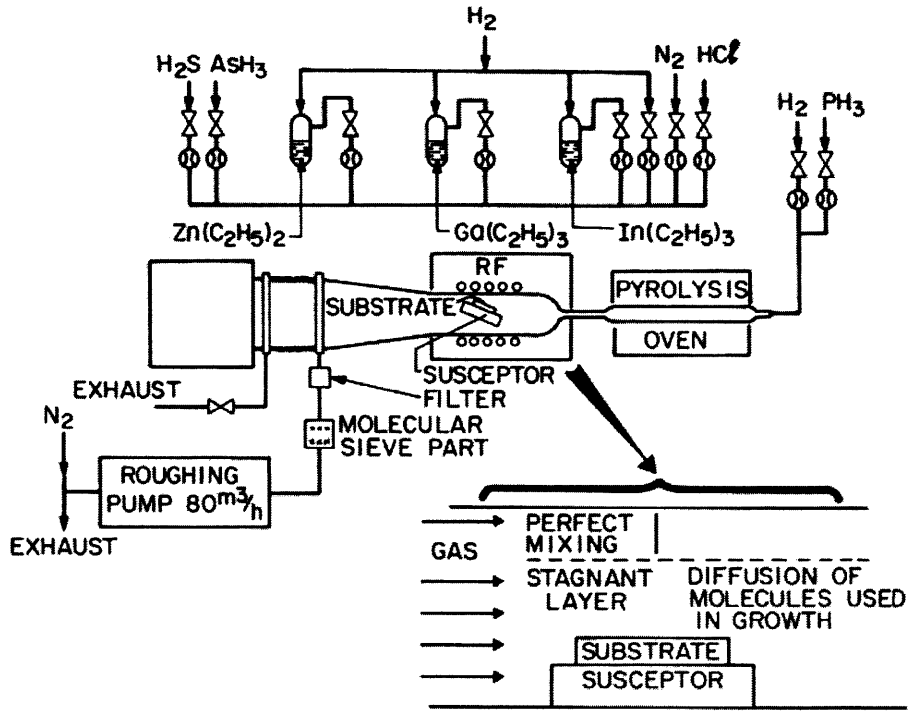

***Figure 2.6** A schematic illustration of an MOVPE system [9].*

In order to grow a typical $Al_xGa_{1-x}As$ heterostructure the GaAs substrate is placed in a tube through which hydrogen gas is passed. The hydrogen acts as a carrier gas to which carefully controlled amounts of gaseous trimethyl gallium, $Ga(CH_3)_3$, trimethyl aluminium, $Al_2(CH_3)_6$ and arsine (AsH_3) are added. These gases act as the sources for the gallium, aluminium and arsenic respectively. Epitaxial deposition of the compounds at the substrate surface occurs as a result of pyrolysis, and a layer of single crystal of $Al_xGa_{1-x}As$ is formed via the following reaction:

$$\tfrac{1}{2}x\mathrm{Al_2(CH_3)_6(g)} + (1-x)\mathrm{Ga(CH_3)_3(g)} + \mathrm{AsH(g)} \xrightarrow{\mathrm{H_2(g)\ 600\text{-}750^\circ C}} \mathrm{Al_xGa_{1-x}As(s)} + 3\mathrm{CH_4(g)}$$

The exact aluminium mole fraction in the $Al_xGa_{1-x}As$ may be altered by varying the partial pressures of the gases involved. Hence, the composition of the deposited layer may be accurately controlled. Gaseous sources of dopants, such as silane (SiH_4) and diethyl zinc $Zn(CH_3)_6$, may also be introduced, thereby allowing the production of *n* or *p*-type materials.

Typically, growth takes place at atmospheric pressure at a temperature of 650°C [8]. Under these conditions, the $Al_xGa_{1-x}As$ growth rate can be as low as 0.5 nms^{-1}. Accurate control over the source gasses allows modern MOVPE equipment to change the gas composition over time-scales of the order of 0.1 s. The time required for a change in the gas composition is therefore much less than the time required for the deposition of a single monolayer. This means that very abrupt changes in material composition may be achieved, permitting the growth of well-defined super-lattice structures with reproducible control over composition, thickness and doping on an atomic scale.

2.6 Device Substrates

The two most widely used substrate materials for III-V devices are GaAs and InP. Each material is produced as a large cylindrical single crystal ingot. This is then sawn up to produce the individual substrate wafers. These wafers will have preferential cleavage planes that will be of great use once the devices have been grown on the substrate (e.g. to allow individual devices to be obtained from the wafer by cleaving). It is important then, to first determine the orientation of the crystal planes within the ingot. This is achieved by X-ray diffraction, and a flat is then ground along the length of the ingot, aligned to a (110) plane of the crystal. This is referred to as the *major flat*. An additional shallower flat may also be added. This is referred to as the *minor flat* and its position relative to the major flat indicates the doping type and crystal orientation. An individual wafer substrate is illustrated in

fig 2.7. There are several internationally recognised conventions governing the standard orientation of flats or the thickness and diameter of substrates. Typical diameters include 2", 3" and 4" although wafers as large as 6" are in standard use for GaAs substrates. A typical 4" GaAs wafer has a thickness of 600 μm, with the thickness varying proportionately with the wafer diameter. InP, being a more brittle material, requires a comparatively thicker substrate for each diameter of wafer.

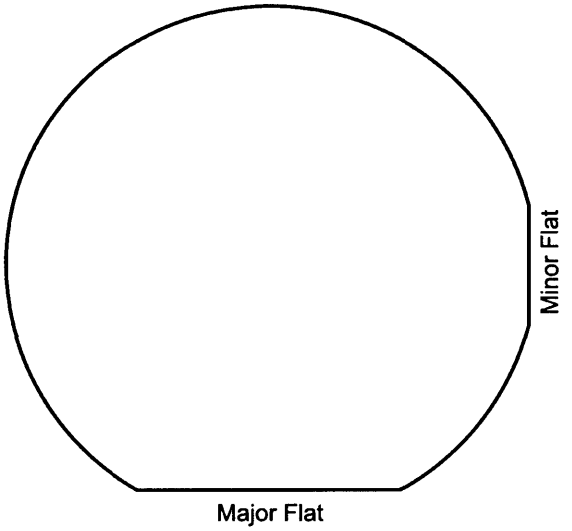

Figure 2.7 *A typical substrate wafer, illustrating the position of the major and minor flats.*

When slicing up the wafer ingot, it is not always desirable to saw it directly along a (100) axis. A misorientation angle, α, is sometimes introduced, typically 2° or 10° towards the (110) axis. As illustrated in fig 2.8, this produces steps of atomic planes along the surface of the substrate. This can greatly aid in the epitaxial growth of layers on the substrate by providing nucleation sites for atoms.

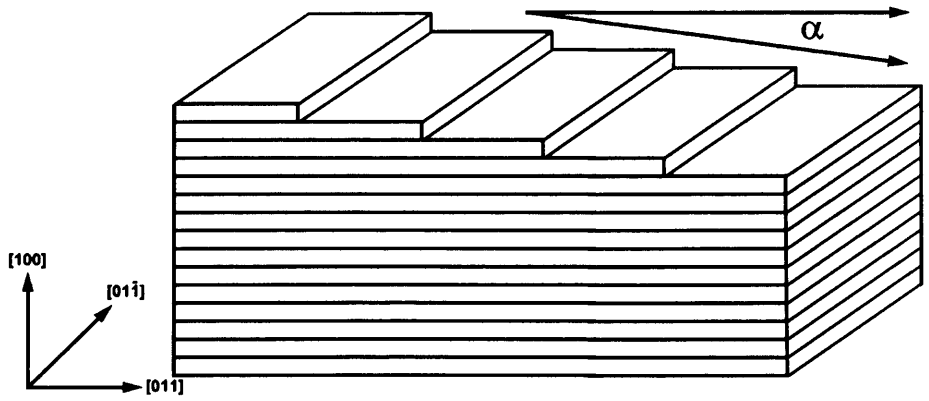

Figure 2.8 Atomic steps on the surface of a substrate wafer produced by misorientated cutting of the ingot at an angle α from the (100) axis. Crystallographic directions are indicated in the diagram.

2.7 The pn Junction

A *pn* junction is formed in a semiconductor material at the boundary between *n* and *p*-type regions. Electrons and holes can recombine at the junction to produce light. This property is exploited in optoelectronic devices, such as the light emitting diode (LED) and semiconductor laser diode (SLD or LD).

Practical *pn* junctions are fabricated from single crystals in order to minimise defects and impurities, increasing the chance of electron-hole recombination and ensuring good electrical conductivity. Traditionally, diffusion doping methods were employed, whereby the dopant diffuses into the crystal from the surface under heating. The drawback of this method was that it produced ill-defined junctions. Nowadays, several techniques are available. Ion implantation, for example, involves bombarding the starting material, *n*-type say, with the required species of ions. This produces sharper junctions, but damages the crystal lattice by increasing the number of dislocations and interstitial atoms. Another method is to deposit doped crystalline films using one of the epitaxial techniques outlined above. Metallurgically speaking, the growth of well-defined junctions is possible. However, a junction can never be perfectly abrupt electrically. It always extends over a finite distance due to the diffusion of carriers.

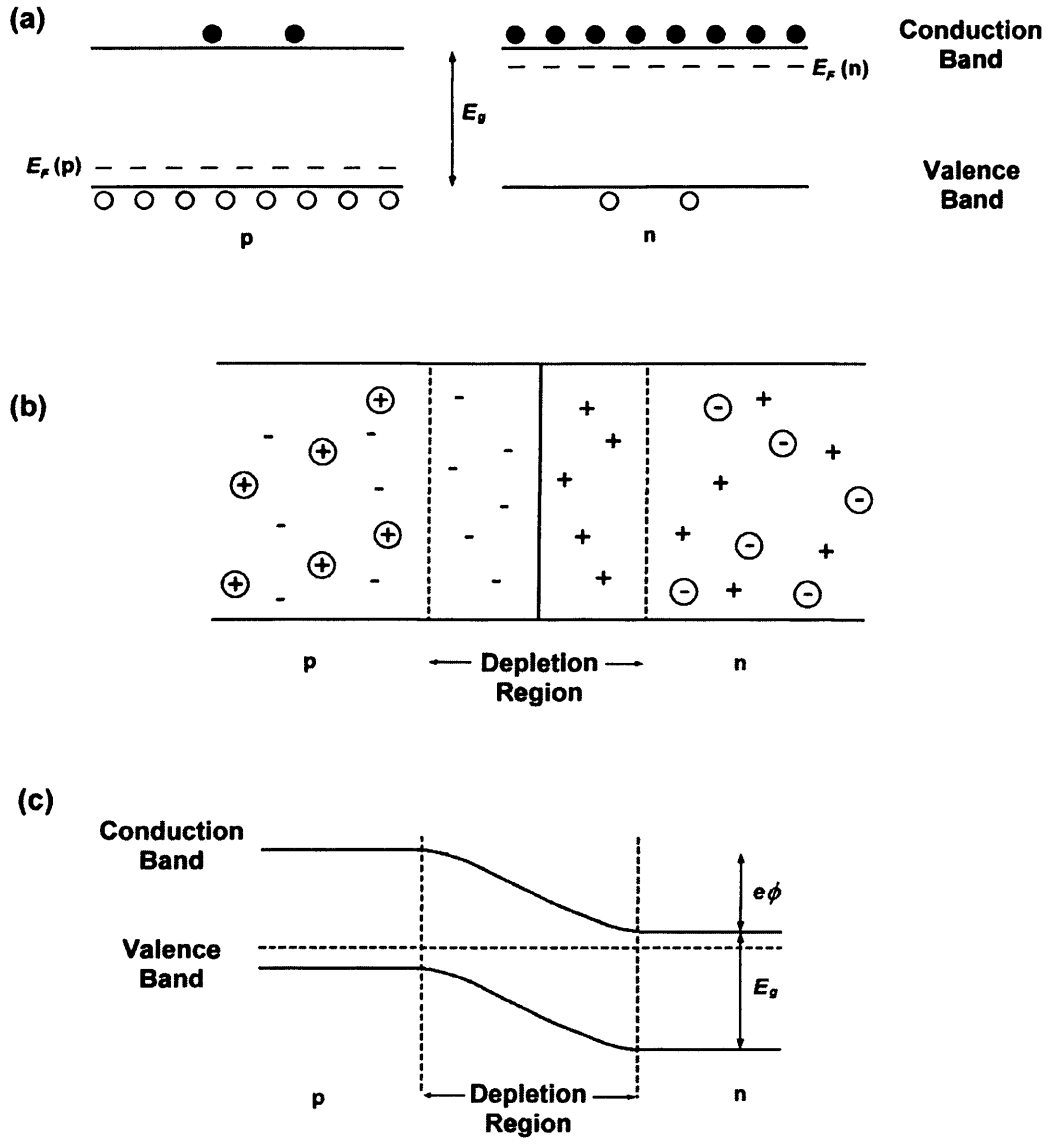

Figure 2.9 (a) The Fermi energy is closer to the valence band edge in a p-type semiconductor and to the conduction band edge in an n-type semiconductor. Conduction electrons and holes are represented by solid and open circles respectively. (b) When the p- and n-type regions are in contact, diffusion of carriers results in a depletion region in which there are virtually no carriers. (c) The distribution of charge at a pn junction provides a contact potential ϕ across the junction so that the conduction band edge on the p-type side of the junction is at an energy $e\phi$ higher than that on the n-type side [7].

Despite the fact that *pn* junctions are fabricated from single crystals, it is useful when trying to understand their properties to imagine they are formed by the joining together of two separate pieces of crystal, one doped *p*-type, the other doped *n*-type. The situation is illustrated in fig 2.9. As separate entities, the two pieces of

crystal are electrically neutral, the charge of the excess free carriers being balanced by the residual charge on the ionised donor atoms from which they arise. When they are brought together, differences in electron and hole concentrations between the two layers causes a diffusion of charge carriers to occur at the junction. If an electron enters the *p*-type material from the *n*-type, it is unlikely that it will travel far before recombining with a hole. The same can be said of any holes entering the *n*-type material from the *p*-type. It can be seen that this flow of carriers produces an immediate loss of charge neutrality in the region either side of the junction (it should be noted however, that the bulk materials are still neutral). The *n*-type material will become positive because of a deficiency of electrons that have recombined with holes, and the *p*-type material will become negative in a similar manner. The resultant potential gradient across the junction forms a dipole layer that brings the Fermi levels of the two materials into line. Thus a dynamic equilibrium is established that acts as a barrier to any further charge movement. However, this equilibrium may be changed by the application of an external current, which will allow steady diffusion currents to occur once again, resulting in carrier injection across the junction.

The resulting charged layer either side of the junction is referred to as a *depletion region* and its thickness depends of the carrier concentrations of the two materials involved. Higher carrier concentrations produce thinner depletion regions. Typical thicknesses are of the order of about 0.05 μm for carrier concentrations of $n = 10^{18}$ cm^{-3} [5].

2.8 Edge Emitting Lasers

The most basic edge emitting lasers consist of heavily doped (~10^{18} cm^{-3} [9]) homogeneous *pn* junctions (fig 2.10a). On the application of a sufficient forward bias, carriers flow into the depletion region where electron-hole recombination occurs. The edges of the laser are formed by utilising the natural cleavage planes of the semiconductor material. If carefully cleaved, the edges produce very flat facets that act as mirrors. The reflectivity of such mirrors is up to 30 %, so while some of the light escapes, a proportion is reflected back into the laser cavity and hence the junction. Since the energy of this light is the same as the bandgap, stimulated

emission occurs. Typical dimensions for such a device are 300 μm × 100 μm × 200 μm [12], with a threshold current (i.e. the current at which stimulated emission occurs) of 10 mA [12]. Below threshold, the laser acts as a conventional light emitting diode [13].

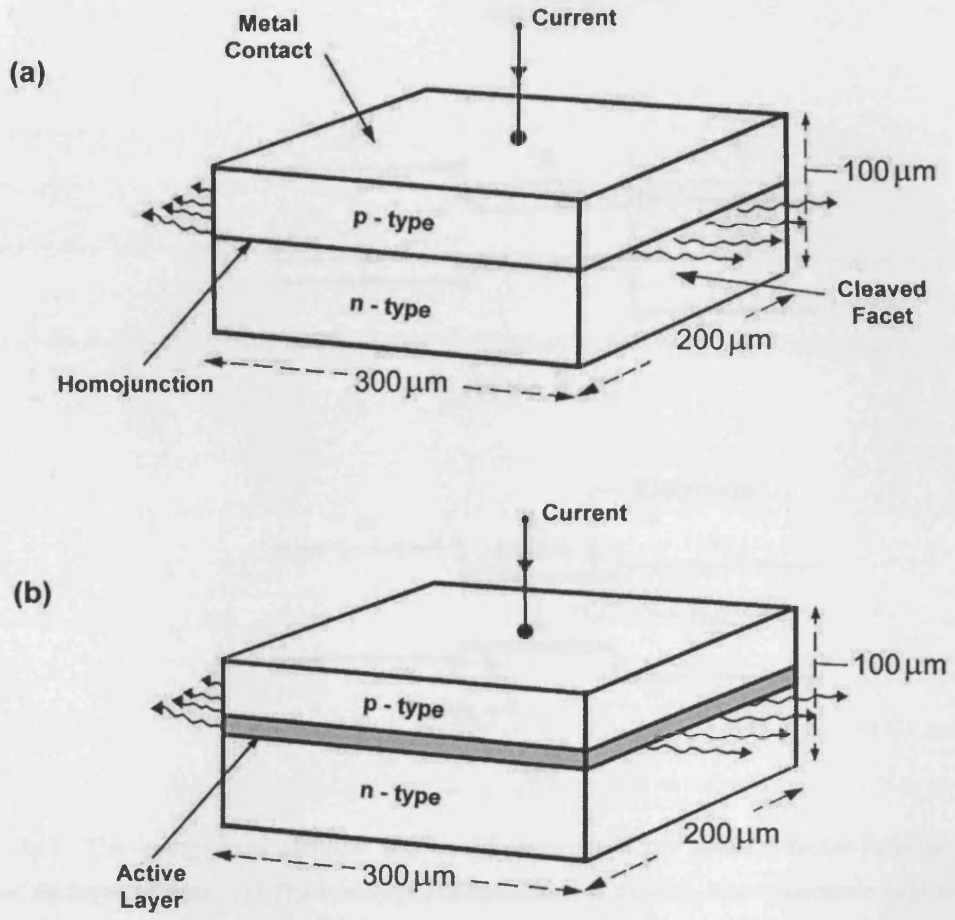

Figure 2.10 *Schematic illustration of* ***(a)*** *homostructure and* ***(b)*** *double heterostructure semiconductor lasers with their typical physical dimensions. The shaded area shows the thin (~ 0.2 μm) active layer of a semiconductor material whose bandgap is slightly lower than that of the surrounding cladding layer [14].*

The design described above is inefficient however, as the nature of the *pn* junction means that there is no boundary to the depletion region in which carrier recombination can occur. Subsequently, carriers can be lost to diffusion before recombination occurs. As a result, these so-called *homojunction* lasers have relatively high threshold currents.

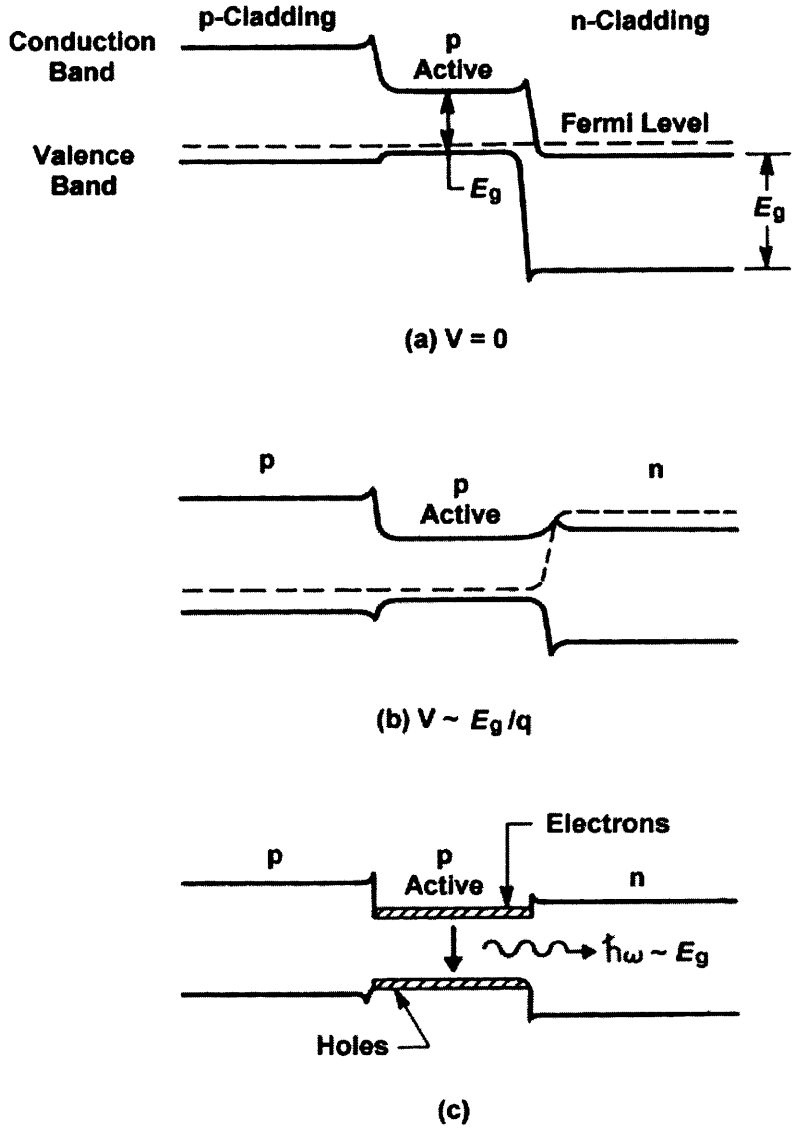

Figure 2.11 The energy-band diagram of a double-heterostructure semiconductor laser at (a) zero bias and (b) forward bias. (c) The bandgap discontinuities at the two heterojunctions help to confine electrons and holes inside the active region, where they recombine to produce light [14].

In order to produce a laser with a lower threshold current (and hence a smaller power consumption), we must employ a double heterostructure design (fig 2.10b). This is formed by sandwiching a thin layer (~ 0.2 μm [15]) of lattice matched material of smaller bandgap and higher refractive index between two oppositely doped semiconductors. Here, carriers are confined to the depletion region by a potential barrier on either side (as illustrated in the band diagrams of fig 2.11), establishing higher electron and hole densities and allowing more efficient recombination when a forward bias is applied [16, 17]. A further advantage of this design is that optical

feedback is much improved due the difference in refractive index between the active region and cladding layers.

An essential feature of the active region is the *quantum well*, a thin layer (typically less than 10 nm [6]) of semiconductor material with a smaller bandgap than the surrounding cladding layers. Such thicknesses are close to the De-Broglie Wavelength of the holes, and quantum effects become apparent. Transitions now occur between discrete energy levels, thereby decreasing the wavelength distribution of the emitted light. Despite being mostly governed by the choice of material, the quantum well also allows a certain degree of control over the wavelength of the emitted light. Careful consideration of the well width leads to adjustments in the energy levels within the well.

The power of the light emitted is also a concern, and generally, the power obtained from a single well is quite low. It is therefore desirable to employ multiple quantum wells, separated by a cladding material.

Both homojunction and heterojunction lasers emit light in the plane of the junction. They are therefore called *edge emitting* lasers. The designs outlined above are relatively simple. More complicated lasers with various different material layers and types of heterojunctions have been utilised to lower the threshold and increase the gain of the laser. Despite all this however, they are considered too large and too power consuming for many purposes. Another problem is that they are edge emitting. Ideally, lasers should emit light in a direction normal to the plane of the junction, allowing the emission of light over larger areas. Research into the design of these so-called *surface emitting* lasers has resulted in the development of one of the most important optoelectronic devices, the *vertical cavity surface emitting laser* (VCSEL). The design here is somewhat different to conventional laser diodes, and shall be discussed in the next section.

2.9 The Vertical Cavity Surface Emitting Laser (VCSEL)

2.9.1 Introduction

As outlined in the previous section, the vertical cavity surface emitting laser has many advantages over its edge-emitting counterpart. Devices may be manufactured and tested on a single wafer. In addition, large arrays of devices can be created, allowing the possibility of optical interconnects and optical neural networks [6]. VCSELs are also typically smaller than their edge emitting equivalents, and have smaller output wavelength distributions, making them particularly useful in the telecommunications industry [5].

The first VCSEL device is attributed to Soda, Inga *et al* in 1979 at the Tokyo Institute of Technology [6]. The first devices used a double heterostructure design with an InGaAsP active layer, which emitted light at 1.3 μm [6]. Since then, the basic VCSEL design, as illustrated in fig 2.12 has changed very little.

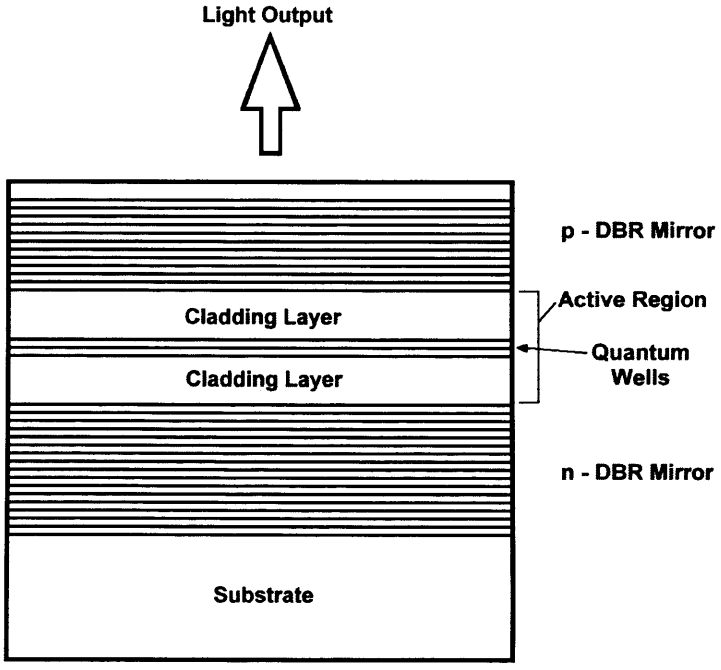

Figure 2.12 A schematic diagram of a vertical cavity surface emitting laser.

The cavity length of a VCSEL is very short, typically between 1 and 3 wavelengths of the emitted light. As a result, the optical feedback that would be produced by cleaved facets is not efficient enough to trigger stimulated emission. This means that very reflective mirrors must be used. Modern VCSEL devices utilise *distributed Bragg reflectors* (DBRs) that have reflectivities greater than 99.9 %, allowing low threshold currents.

2.9.2 Distributed Bragg Reflectors (DBRs)

The distributed Bragg reflectors (known also as *mirror stacks* or *Bragg stacks*) are an integral part of the VCSEL, since they allow for the emission of light in a direction perpendicular to the active region. Placed above and below the active region, the DBRs are formed by laying down alternating layers of semiconductor materials with a difference in refractive index. Each of the layers has a thickness of ¼ of the output light wavelength in the material (typically 60 nm [5]), and the superlattice lowers the threshold current of the VSCEL by producing a large amount of optical feedback into the active region. Indeed, reflectivities of greater than 99.9 % may be produced by appropriate consideration of the design of a DBR.

The efficiency of the DBR depends on the difference in refractive index between the alternating layers. Interfaces with a higher difference in refractive index achieve a higher reflectivity, as do a large number of reflecting layers. Traditionally, DRBs were fabricated using GaAsInP/InP. The drawback here however, is the low difference between the refractive indicies of the two materials, meaning that a large number of periods are required. Since the DBRs also carry current in the device, the large number of layers increases the resistance of the VCSEL. Recently, much interest has been shown in $Al_xGa_{1-x}As/Al_yGa_{1-y}As$ ($x \neq y$) DBRs. Excellent lattice matching and higher differences in refractive index ($n_{GaAs} = 3.65$, $n_{AlAs} = 3.18$) mean that the DBRs are easier to manufacture and require a smaller number of layers in order to achieve the high reflectivities demanded in VCSEL devices.

One disadvantage of using materials with large refractive index differences is that potential barriers are created at the heterojunction interfaces due to bandgap differences. This problem can be alleviated by using graded heterobarrier interfaces that maintain the refractive index differences and hence do not reduce the reflectivity

of the DBR. By introducing grades into the superlattice, the composition changes from one layer to the next over a finite distance, in effect, 'smoothing out' any changes in potential barriers. For example, within an $Al_xGa_{1-x}As/Al_yGa_{1-y}As$ superlattice, the mole fraction of Al is varied slowly from the value x to the value y in the grades, thereby allowing easier electron transport. Fig 2.13 illustrates differences in the variation in potential between a graded and ungraded superlattice.

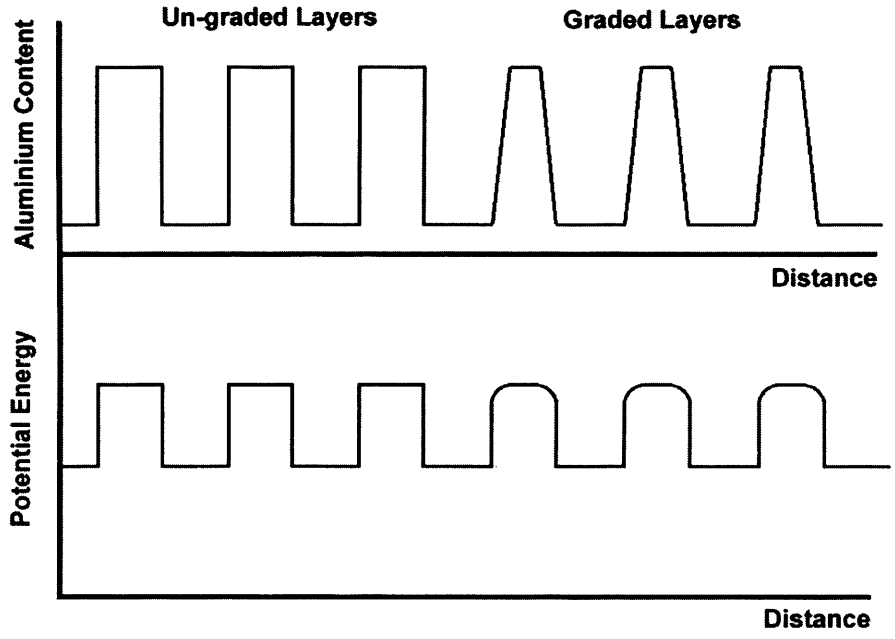

Figure 2.13 An illustration of the variation in potential between graded and ungraded superlattice structures.

2.9.3 Current Confinement

Other than modifying the properties of the DBRs, there are a number of methods of reducing the threshold current of a VCSEL device. The most effective method is to limit the cross-sectional area in which gain occurs. There are a number of different ways that this may be achieved. The simplest method is to etch a pillar down to the active layer (fig 2.14a). The added advantage here is that the large difference in refractive index between the air and device material acts to guide the light emitted. Problems with this type of structure are the loss of carriers due to

surface recombination at the sidewalls and poor dissipation of heat from the laser cavity.

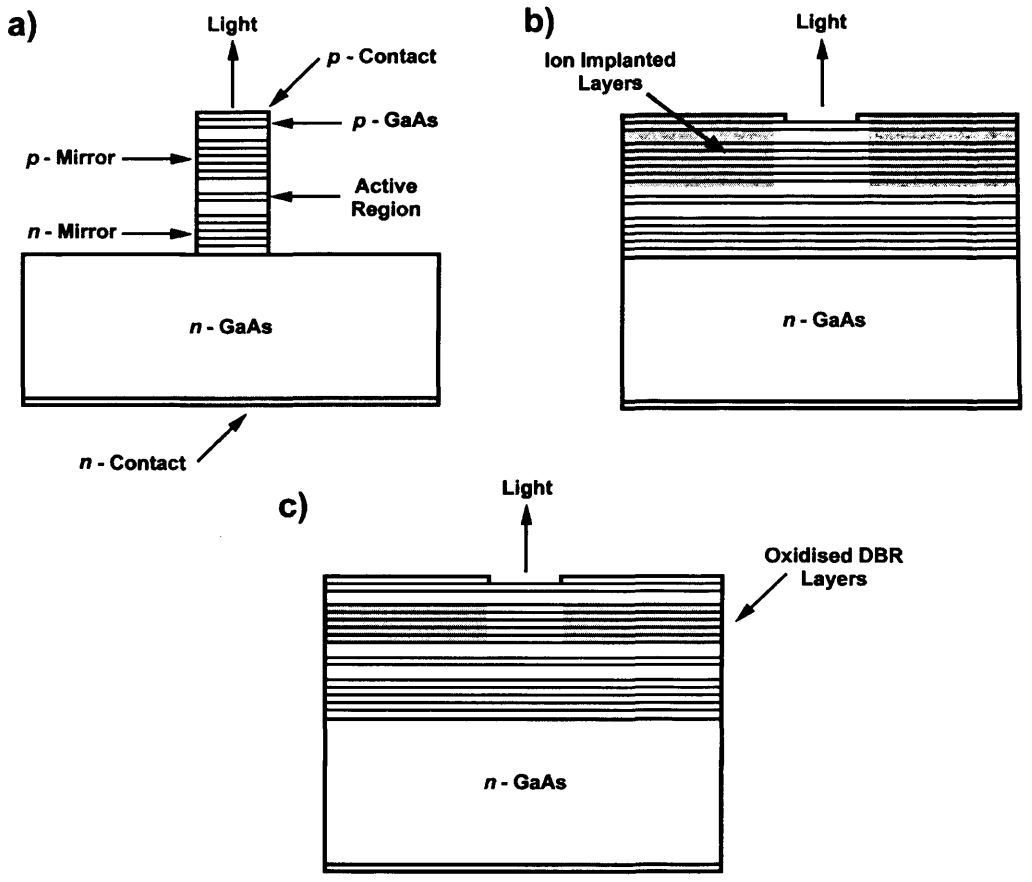

Figure 2.14 *Current confinement methods* ***(a)*** *Etched pillar;* ***(b)*** *Ion implantation (grey shading indicates the ion implanted regions);* ***(c)*** *Oxide confinement (grey indicates sections of the layers that have been oxidised) [14].*

A more complicated way of confining the current is to preferentially damage areas of the upper mirror stack by selectively implanting ions into the semiconductor material (fig 2.14b). By bombarding selected areas of the device with protons, they may be turned into insulating material. A small opening is left undamaged at the centre of the device, and current flow (and hence laser emission) occurs in this region only. This method, known as *ion implantation* does have its drawbacks, however. The bombarding of the semiconductor with ions tends to damage the crystal structure of the implanted material and so it must be used with caution within close proximity of the active layer.

The last method achieves current confinement by using laterally grown thermal oxide layers above, and or below the active region, as shown in fig 2.14c. In a steam environment of 350°C to 500°C the $Al_xGa_{1-x}As$ material in the Bragg reflectors is converted into an oxide layer. The rate of formation of the oxide layer increases with the Al content of the material; thus oxide forms fastest in those layers with the highest aluminium content.

Apart from the fact that a natural property of the material is utilised, there is one other main advantage. The insulating properties of the oxide result in a reduction of carrier leakage currents into the non-active region.

Due to its obvious usefulness as a current confinement method, much research has been devoted to the lateral oxidation of devices. An understanding of the formation of III-V oxides is also of pivotal importance to the research discussed within this thesis. Chapter 3 is devoted to the mechanism of oxidation in III-V compounds and it is not discussed further in this chapter.

2.10 References

[1] R N Hall, G E Fenner, J D Kingsley, T J Soltys & R O Carlson. *Physics Review Letters*. **9**, 366 (1962).

[2] M I Nathan, W P Dumke, G Burns, F H Dill Jr. & G Lasher. *Applied Physics Letters*. **1**, 62 (1962).

[3] T M Quist, R H Rediker, R J Keyes, W E Krag, B Lax, A L McWhorter & H J Zeiger. *Applied Physics Letters*. **1**, 91 (1962).

[4] N Holonyak Jr. & S F Bevacqua. *Applied Physics Letters*. **1**, 82 (1962).

[5] S M Sze. *Physics of Semiconductor Devices (2nd Ed.)*. John Wiley and Sons (1981).

[6] J Singh. *Semiconductor Devices - An Introduction*. McGraw Hill International Editions (1994).

[7] R Turton. *The Physics of Solids*. Oxford University Press (2000).

[8] D V Morgan & K Board. *An Introduction to Semiconductor Microtechnology*. John Wiley and Sons (1996).

[9] S D Smith. *Optoelectronic Devices*. Prentice Hall International Series in Optoelectronics (1995).

[10] A Y Cho. *Journal of Vacuum Science and Technology*. **16**, 275 (1979).

[11] R D Dupuis. *Journal of Crystal Growth*. **55**, 213 (1981)

[12] J J Liou. *Advanced Semiconductor Device Physics and Modelling*. Artec House Publishing (1994).

[13] H P Myers. *Introductory Solid State Physics (2^{nd} Ed.)*. Taylor and Francis Publishing (1997).

[14] G P Agrawal & N K Dutta. *Semiconductor Lasers (2^{nd} Ed.)*. Van Nostrand Reinhold (1993).

[15] M Ettenberg. *Applied Physics Letters*. **27**, 652 (1975).

[16] H Kroemer. *Proceeds of the IEEE*. **51**, 1782 (1963).

[17] Zh I Alferov & R F Kazarinov. Authors certificate 181737 (USSR) (1963).

Chapter 3

The Oxidation of III-V Semiconductor Materials

3.1 Introduction

Much research has been dedicated to understanding the oxidation of III-V materials, particularly GaAs. Despite this, there is still no generally accepted model of the detailed oxidation processes. However, it is possible to appeal to the mechanism by which silicon oxidises for guidance, as this is well understood. Most research suggests that III-Vs do indeed have a superficially similar form of oxidation to that of silicon. Unfortunately, it is complicated by the competition that arises between different oxidising species. For example, one element may diffuse, evaporate, dissolve or oxidise more quickly. This suggests that the chemical composition of the oxide in the bulk or at its interface may be non-uniform.

Early studies of the (110) surfaces of semiconductors employed electron-beam-excited *Auger electron spectroscopy* (AES) in order to determine the oxygen uptake at room temperature and under vacuum [1]. The experiments concentrated mainly on the bulk oxidation of GaAs, InAs and InP. More recently, there has been great interest shown in the oxidation of AlGaAs, particularly in relation to the production of VCSEL based mesa devices. However, experiments here have concentrated on oxidation of a thin encapsulated layer. Extensive lateral (sub-surface) oxidation of the heterostructures is carried out in a steam atmosphere at relatively high temperature to allow the production of current apertures (see section 2.9.3).

By contrast, this research is concerned with the room temperature oxidation of the (110) surfaces of III-V heterostructure devices. The material in the thin oxidising layers might be considered to be partially bound by other layers and may therefore not oxidise in the same way that either bulk or unbound material would.

A summary of the results of past oxidation studies is outlined in this chapter. An attempt is then made to present a coherent view of how cleaved III-V heterostructures might oxidise at room temperature.

The chapter begins with a discussion of the oxidation regimes associated with III-V materials. A review of some of the factors that affect the oxidation of III-V materials is then presented. Finally, conclusions are drawn as to how the oxidation of bound III-V compounds occurs, and a summary of expected observations is presented.

3.2 Oxidation Regimes

There are several well-defined processes by which the oxidation of III-V materials may occur. Individually, they are reasonably well understood, but their relationship and the conditions for which a particular process is the most favourable is not. Indeed, in the case of bound layers, there are often conflicting reports within the literature as to in what layer thickness and temperature regime each particular mechanism dominates.

After a crystal is cleaved, an initial protective oxide layer is formed by *activated adsorption.* This is an extremely rapid process, occurring virtually instantaneously. A strong electric field is then set up across the oxide layer that allows the oxidation to proceed further by pulling substrate or oxygen ions through the non-reacting oxide layer. Once the oxide reaches a certain thickness, the electric field is no longer strong enough to aid oxidation, and the oxidation becomes either diffusion or reaction rate limited. However, in the case of elevated temperatures, the effects of the latter two mechanisms may be evident immediately, and the effects of field aided oxidation may not contribute significantly to the extent of oxidation.

In the discussion that follows, each of these mechanisms shall be dealt with, indicating the most favourable conditions under which they occur.

3.2.1 Activated Adsorption Regime

When any freshly cleaved material is exposed to air, an initial adsorption layer is formed quickly on its surface. For III-V materials, an exposure time of 0.1 ms is required to produce a monolayer film covering the surface [2]. Molecules of the oxidising species (e.g. O_2, H_2O, OH^-, and CO) form bonds with the surface

(substrate) atoms of the material undergoing oxidation. These bond strengths may vary from ~ 0.1 to 3 eV [1]. Weak bonding, where the molecules merely rest on the surface, is referred to as *physisorption*. Stronger bonding involves the transfer of charge between the surface and the adsorbing molecule and is referred to as *chemisorption*. In the case of GaAs, the layer initially consists of Ga_2O_3 together with elemental As [3].

This initial monolayer of growth is uniform across the surface of the semiconductor [1], and acts as a diffusion barrier, resisting further oxidation. If the oxidation is to continue, movement of the reactant species through the protective layer must occur.

At elevated temperatures, there is sufficient energy for the thermal diffusion of the reactant species to be the dominant mode of oxidation. However, at room temperature, where thermal diffusion is less effective, *field-aided transport* could become the dominant mechanism.

3.2.2 Field-aided Regime

Field-aided oxidation was originally proposed to explain the oxidation of aluminium, and is known as the *Cabrera-Mott mechanism* [2]. It has also been reported to be the main uptake mechanism for the (110) surfaces of GaAs, InAs, InP [4] and $Al_xGa_{1-x}As$ (up to $x = 0.7$) [5] at room temperature, and has subsequently been proposed to be effective for all III-V semiconductors [4].

After the initial oxide film is laid down, a strong electric field is induced by the contact potential difference between the substrate and adsorbed oxygen. This field allows ions to move through the non-reacting oxide layer without the assistance of temperature.

A theoretical analysis yields an inverse logarithmic growth law for oxide thickness h, as a function of time, t. This may be written as:

$$\frac{1}{h} = \frac{1}{h_0} + \frac{1}{h_1}\ln(t/\tau) \qquad (3.1)$$

where

$$h_1 \propto e^{(eV/k_BT)}$$

Here, h_0 is the oxide film thickness at which field-aided growth begins and τ is a time scaling factor. The term h_1 is proportional to the potential V (known as the *Mott potential*) across the oxide layer. k_B is the Boltzman constant. Although this potential difference may remain constant, the electric field strength diminishes as the oxide layer thickens. This mechanism is consequently only effective for relatively thin films [5]. As the electric field weakens, thermal diffusion becomes the dominant transport mode, and oxidation becomes either reaction rate or diffusion limited, as discussed in the next section.

3.2.3 Diffusion and Reaction Rate Limited Regimes

3.2.3.1 The Model of Silicon Oxidation

Extensive research on silicon [6] has shown that in order for thermal oxidation to take place, oxygen must enter the oxidising layer, be transported to the oxidation front and react with the crystal substrate to form the oxide. Clearly, the slowest of these processes will limit the oxidation rate.

Diffusion of the species through the oxide is governed by *Fick's first law*,

$$J_F = -D\left(\frac{\partial N}{\partial h}\right) \tag{3.2}$$

which states that the number of diffusing atoms per unit area (the flux), J_F, is proportional to the concentration gradient, $\partial N/\partial h$, of the oxidant in the oxide. This situation is illustrated in fig 3.1. The constant

$$D = D_0 e^{(-E_A / k_B T)} \tag{3.3}$$

links the temperature of oxidation, T, and the activation energy of diffusion of the oxidant through the existing oxide, E_A. The constant D_0 is the value of D extrapolated to infinite temperature. It is given as,

$$D_0 = \tfrac{1}{3}\lambda \bar{c} \qquad (3.4)$$

where λ is the mean free path of the diffusing particles, and $\bar{c}$ is their mean speed. Since the mean speed is dependent on temperature, then so too is the value of D_0.

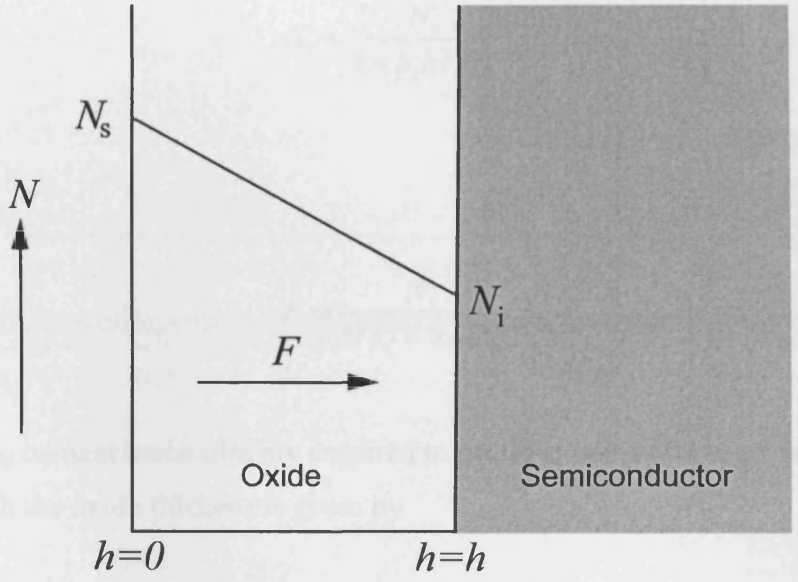

Fig 3.1 *The oxidation of a semiconductor material. An oxide of thickness h forms on the surface. Oxidant molecules of flux F impinge on the oxide-semiconductor interface i. Superimposed is a plot of the oxidant concentration, N. This concentration varies linearly through the oxide, having a maximum value N_s at the oxide surface and a minimum value N_i at the oxidation front.*

For the case of bulk material, consider a unit-area cross section of semiconductor plus oxide, in which the oxidant concentration N (molecules/m^3) varies linearly from N_s at the surface to N_i at the interface, as shown in fig 3.1. If the flux F of oxidant molecules passing through the oxide satisfies Fick's first law of diffusion, then

$$F = -D\frac{\partial N}{\partial h} = \frac{D(N_s - N_i)}{h} \qquad (3.5)$$

If the flux F_i represents the rate at which the oxidant molecules are consumed at the interface in the production of oxide, it is reasonable to assume that $F_i \propto N_i$:

$$F_i = k_i N_i \tag{3.6}$$

In a steady state, the two fluxes must be equal. Equating equations 3.5 and 3.6 obtains

$$N_i = \frac{N_s}{1 + k_i h / D_0} \tag{3.7}$$

and

$$F = \frac{N_s}{1/k_i + h/D_0} \tag{3.8}$$

If N_m oxidant molecules are required to produce unit volume of oxide, then the rate at which the oxide thickens is given by

$$\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{F}{N_m} = \frac{N_s / N_m}{1/k_i + h/D_0} \tag{3.9}$$

Integrating this from 0 to t gives a quadratic in h that may be solved to give

$$h = \frac{A}{2}\left[\sqrt{\left(1 + \frac{t + t_0}{A^2/4B}\right)} - 1\right] \tag{3.10}$$

where t_0 is a characteristic time-scale. This may be written more simply as

$$h^2 + Ah = Bt \tag{3.11}$$

where $A = 2D_0/k_i$ and $B = 2D_0N_s/N_m$.

In the case of large oxide thicknesses, equation 3.11 reduces to

$$h \approx (Bt)^{1/2} \tag{3.12}$$

Since the constant B is proportional to the diffusion coefficient D_0, the oxidation in the limit of thick oxides is diffusion rate limited, and has a parabolic dependence.

In the case of small oxide thicknesses, equation 3.11 reduces to

$$h \approx \frac{B}{A}t \tag{3.13}$$

Here, the constant B/A is not dependent on the diffusion coefficient. The oxidation in this case is assumed to be reaction rate limited, and has a linear dependence.

These dependencies make sense from a physical point of view. In the case of diffusion rate limited oxidation, it is expected that the oxidation rate will decrease with increasing oxide thickness, as the oxidants have a larger distance to travel to reach the oxidation front. In the case of reaction rate limited oxidation, a decrease in the oxidation rate would not be expected, as the rate does not depend on how much oxide the reactants have to pass through.

3.2.3.2 Application of the Model to III-V Materials

To a certain extent, the model of silicon oxidation outlined above has been successfully applied to explain the oxidation of III-V materials in the bulk case. However, the effects of several different species of element must be taken into consideration. There will be more than one oxidation reaction occurring, the slowest of which determines the oxidation rate in the rate limited case. Also, the different mobilities of the species within the oxide will clearly have an effect on the reactions. Added to this is the fact that the group V element can, at higher temperatures, evaporate from the oxide [7]. These factors are likely to result in a far more complicated oxidation mechanism.

Studies on bulk GaAs, AlGaAs and InP carried out at room temperature indicate no evidence of a linear reaction rate dependence [8]. The oxidation in these cases was found to be diffusion rate limited. Indeed, research has shown that samples stored in air for as long as eight years continue to fit the parabolic curve [1]. At elevated temperatures, the oxidation conforms to that of the silicon model: it is reaction rate limited for thin oxides and diffusion rate limited for thick oxides [9,10].

However, in cases where the oxidising material is bound, there may be further deviations from the silicon oxidation model. Indeed, the oxidation mechanisms are not particularly well understood in this case, despite much research.

Studies on the lateral subsurface oxidation of bound AlGaAs layers in VCSEL structures at high temperatures and in the presence of steam show that the dominant limiting process is dependent on temperature, aluminium concentration and layer thickness [9,10]. These show that reaction rate limited oxidation is favoured by decreases in temperature, aluminium concentration and layer thickness. Conversely, diffusion rate limited oxidation is favoured by increases in temperature, aluminium concentration and layer thickness. However, the wide range of temperatures, layer thicknesses and aluminium concentrations investigated in the literature make it difficult to directly compare them and quantify under exactly what conditions the transition between reaction rate and diffusion limited oxidation occurs.

Research into $Al_xGa_{1-x}As/GaAs$ superlattice structures oxidising at room temperature [11] has claimed direct observation of the field-aided regime followed by a transition to diffusion limited oxidation. A series of $Al_xGa_{1-x}As$ layers were grown, separated by GaAs for various compositions $x = 0.12 - 0.72$. The effects of layer thickness on oxidation rate were not taken into consideration. Layers of random thickness were grown, increasing from 100 nm ($Al_{0.12}Ga_{0.88}As$) to 250 nm ($Al_{0.72}Ga_{0.28}As$).

To account for the oxidation of GaAs in the sample, the heights of oxidised GaAs were added to the mean height difference between $Al_xGa_{1-x}As$ and GaAs. This was achieved by using data obtained from another sample containing a $GaAs/Ga_{0.5}In_{0.5}P$ superlattice. Oxide heights of GaAs were measured with respect to $Ga_{0.5}In_{0.5}P$, which was assumed not to oxidise. The variation of oxide height with time obtained from this study is plotted in fig 3.2.

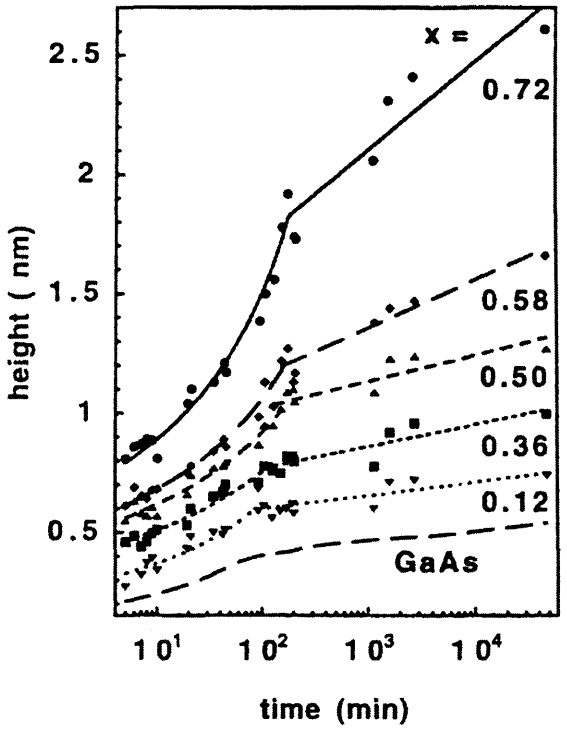

Fig 3.2 Oxide height versus time data for $Al_xGa_{1-x}As$ [11].

It is clear from the data that a change in rate limiting process is likely. The researchers attribute this to a transition between field-aided and diffusion limited oxidation. Transition times showed a dependence on Al mole fraction, ranging between 70 min ($Al_{0.12}Ga_{0.88}As$) and 150 min ($Al_{0.72}Ga_{0.28}As$). Subsequent oxidation is observed to follow a parabolic form, and is claimed to proceed via diffusion limited growth.

3.3 Factors Influencing Oxidation Rate

Careful consideration of the various factors influencing the kinetics of the oxidation rate of III-V materials is clearly warranted. The following discussion summarises the current understanding of these factors. It should be remembered at all times, that current understanding of the oxidation of III-V materials is far from complete. All the factors outlined below are likely to effect the oxidation of the material to some extent.

3.3.1 Width of Semiconductor Layer

In multilayer devices, the width of the oxidising layer has been shown to be of significance, particularly for AlGaAs based structures, where the presence of aluminium leads to more rapid oxidation [11]. It has been determined that the oxidation rate decreases for thinner layers. Two explanations have been proposed to explain this effect.

The first supposes that the number of available oxidation nucleation sites (such as crystal impurities or atomic steps) decreases for narrower layers, leading to a reduced oxidation rate.

The second, more complicated explanation [9], proposes that the effect of layer thickness on oxidation rate is a direct result of interfacial energies. This is similar to capillary action and leads to curvature of the oxidation front within the layer. In other words, this curvature is caused by the surface tension between the advancing oxide and the surrounding crystal, and is thus a measure of the stress exerted on the advancing oxide front. Within this model, the reason that the oxidation rate decreases with decreasing layer thickness is obvious. For thinner layers, the radius of curvature is greater, and hence there is more resistance to the oxidation.

3.3.1.1 Effect of Interfacial Energies on Oxidation

The advancing oxidation front, sandwiched between the two unoxidising layers is analogous to a liquid in a capillary tube. Here, the interface between the oxidised and unoxidised material is similar to the liquid-air interface, while the adjacent unoxidising layers can be likened to the capillary tube itself. Fig 3.3 illustrates the situation. An oxidising layer of material A with thickness d is bound on either side by material B. The oxide O, advances from right to left, with an interface of radius of curvature r. The angle θ_c at which the oxidation front meets the unoxidised material is known as the contact angle.

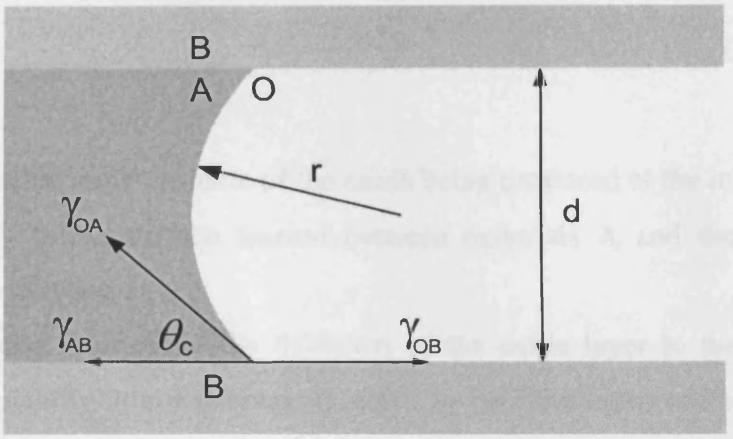

Fig 3.3 *An oxidising material* A, *of thickness d is sandwiched between a non-oxidising material* B. *The oxide* O *consumes material* A *from right to left. The stress at the interface of materials* O *and* A, γ_{OA}, *causes the advancing oxidation front to distort, with a radius of curvature r and contact angle* θ_c. *The two other stresses caused by the interfaces between materials,* γ_{AB} and γ_{OB}, *are indicated on the diagram.*

At equilibrium, forces γ, at the various interfaces occur as a result of interfacial energies. At any instant, these balance each other at the oxidation front thus:

$$\gamma_{AB} + \gamma_{OA}\cos(\theta_c) = \gamma_{OB} \qquad (3.14)$$

Every chemical reaction has an associated energy that determines the driving force of the reaction at a particular temperature T. This is known as the *Gibbs free energy* ΔG^T. A negative value of ΔG^T indicates that that the reaction is energetically favourable and occurs spontaneously.

The stress at the oxidation front acts to introduce an additional term to the Gibbs free energy, thereby resisting the oxidation by increasing the amount of energy required for the reaction to occur. This is known as the *Gibbs-Thompson Effect* [12].

As in the case of a liquid in a capillary, the presence of the contact angle indicates that the interface at the oxidation front opposes the forward movement of the oxide layer. By using thermodynamic relationships, the Gibbs free energy change associated with the interface curvature may be written as

$$\Delta G = \frac{\gamma_{OA} V_O}{r} \tag{3.15}$$

where V_O is the molar volume of the oxide being produced at the interface per unit time, and γ_{OA} is the surface tension between materials A and the oxide O, responsible for the curvature.

The next step is to relate the thickness of the oxide layer to the radius of curvature. This quantity is not necessarily equal to half the layer width. Using a geometric argument, the two can be related via the contact angle by using the following expression:

$$\cos(\theta_c) = \frac{d}{2r} \tag{3.16}$$

Equations 3.14, 3.15 and 3.16 may be combined to obtain the following expression:

$$\Delta G = 2V_O \frac{(\gamma_{OB} - \gamma_{AB})}{d} \tag{3.17}$$

Thus, as the thickness of the oxidising layer is decreased, ΔG increases. It is this term that modifies the activation energy for the oxidation of the layer. Several experiments have been performed with the aim of quantifying the value of ΔG. Typical values of 3.96 eV per nm of curvature [9] have been found for the oxidation of AlAs layers between GaAs.

It can be seen that this is also an oxidation rate limiting process. It adds directly to the reaction rate. Values of the Gibbs free energy for the oxidation reactions are typically a few hundred kJ mol^{-1}. This equates to roughly a few hundred eV being required to oxidise 1 nm^3 of AlGaAs. It can therefore be seen that the effects outlined above can contribute significantly to the Gibbs free energy of the reaction for large curvatures.

As a final observation, it should be noted that the increase in volume of the layers due to oxidation, likely to increase the strain in the surrounding layers, is not

considered by this model. The implications of this increase in volume are discussed later.

3.3.2 Sample Temperature

All studies report that an increased temperature increases the oxidation rate of a sample [1]. This is to be expected, since an increase in thermal energy will aid both diffusion and the rate of the oxidation reaction.

At room temperature, surface oxides of GaAs are known to be relatively smooth, and remain so even when they have been oxidising for a considerable length of time [1] due to the slow oxidation rate. However, at elevated temperatures, thick oxides grow. These oxides are of a very poor quality, as illustrated in fig 3.4. This is one reason that attempts to fabricate electronic devices requiring oxide layers similar to those produced from silicon have met with little success.

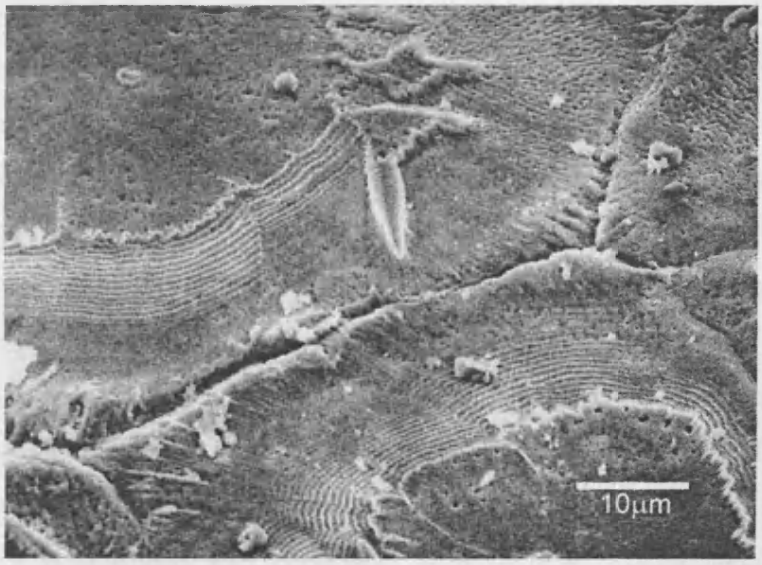

Fig 3.4 An example of poor quality oxide grown on GaAs at elevated temperatures [1].

3.3.3 Chemistry of the Oxides

The chemistry of the III-V oxides is very complicated. Generally, they are a mixture of the oxides of the constituent elements in proportions that reflect the

favourability of their formation. In the case of GaAs, it is a mixture of Ga_2O_3, As_2O_3 and As. In some cases, the oxides can be complex, for example that of InP where $InPO_4$ is formed.

The oxides at the oxidation front may be different, however. At room temperature, the group III element oxide may be found together with the group V element. This is because the III-V compound can quickly reduce the group V oxide. For example, during the oxidation of GaAs, the following reaction occurs at the oxidation front:

$$As_2O_3 + 2GaAs \longrightarrow Ga_2O_3 + 4As \qquad (3.18)$$

The presence of the group V element hinders the oxidation process, and diffuses to the bulk oxide where it oxidises [13]. The group V element may also diffuse to the surface of the oxide and evaporate at high temperatures. The removal of the group V element can also be aided by the presence of hydrogen from moisture in the air as discussed in section 3.3.6.

The process of removal of the group V element has been proved to be ineffective for phosphorus containing compounds. It is thought P has a low diffusion rate and is trapped at the oxidation front where it ultimately forms a complex oxide with the group III element [13].

Table 3.1 gives the oxide compositions at the oxidation front for a variety of III-V materials, at both room and elevated temperatures. It is noted that some thermal oxides have a different composition to those at lower temperatures. This could be due to the fact that the increased thermal energy can allow less energetically favourable reactions to occur, or increase the availability of oxygen.

Ternary compounds will obviously have a more complicated bulk oxide structure. For AlGaAs, it is known to be a combination of Al_2O_3, As_2O_3 and Ga_2O_3. Clearly, the ratio of the group III element oxides is related to the mole fraction of Al. As with the GaAs case, elemental As collects at the oxidation front [14].

For $Al_xGa_{1-x}As$, it has been proposed that at high temperatures (~ 400°C), the transition from linear to parabolic time dependence can be explained by the increased rate of the formation of As_2O_3 of versus its reduction to elemental As [14]. The theory maintains that a steady increasing thickness of the As_2O_3 containing region at

the oxidation front will shift the process from reaction rate to diffusion rate limited. At these temperatures, elemental As may leave the oxidation front more readily. Diffusing species have poor mobility in As_2O_3, explaining the transition to parabolic dependence. It is not known if this effect is compatible with the surface tension model outlined in section 3.3.1.

Table 3.1 *Oxide composition at the oxidation front. Thermal oxides give rise to more compounds at the interface, clearly a result of the elevated temperatures allowing more kinetic barriers to be overcome. Generally, the group V element is present in its natural state at the interface. This is not so for the phosphides. It is thought that this is due to the fact that P cannot diffuse through the oxide layer and is thus trapped at the interface [1].*

III-V Compound	Room Temperature	Thermal Oxide
GaAs	Ga_2O_3 + As	Ga_2O_3 + As
GaSb	Ga_2O_3 + Sb	Ga_2O_3 + Sb
GaP	$GaPO_4$	$GaPO_4$
InAs	In_2O_3 + As	In_2O_3 + As In_2O_3 + As_2O_3
InSb	In_2O_3 + Sb	In_2O_3 + Sb
InP	$InPO_4$	$InPO_4$ + P $InPO_4$

Many of the III-V oxides, particularly that of AlAs, are known to be porous. This occurs particularly when the oxidation takes place in the presence of moisture (see section 3.3.6) and is thought to be related to the removal of the group V element from the oxide [14]. In this case delamination is likely to occur allowing further oxidation of freshly exposed areas. This could have some implications as to the stability of heterostructures that have undergone prolonged oxidation.

3.3.4 Composition of the Oxidising Layer

Since the rate limiting processes will vary between different III-V materials, it naturally follows that each material will have a different oxidation rate. Indeed, the

technique for performing metrology on III-V based semiconductor devices outlined in this thesis exploits this property.

Various reports in the literature [1,2] indicate that the oxidation rates of III-V materials increase in the following order: InP < GaP < InAs < GaAs < AlAs. Phosphide containing materials could be at the bottom of the reactivity order due to the fact that elemental phosphorus accumulates at the oxidation front, reducing the oxidation rate [15].

The effect of Al mole fraction x on the oxidation rate of $Al_xGa_{1-x}As$ has been investigated in recent years [11]. As has been previously discussed in section 2.9.3, a present interest lies in the fabrication of current-confinement apertures for VCSEL devices. These studies, conducted at elevated temperatures, show that the oxidation rate increases with increased Al mole fraction and accelerates when the Al content is above ~ 80 % [16,17]. In the range x = 0.9 - 1.0, the effect of increasing the aluminium content becomes so pronounced that changing the amount by as little as 1 % can lead to significant differences in oxidation rate [18]. A proposed explanation for the increase in oxidation rate with Al mole fraction is the fact that more Al atoms per unit area are available for oxidation at the oxidation front [18], although the behaviour above x = 0.8 indicates that other factors must also be important, for example, the porosity of the oxide.

3.3.5 Oxidation Heights and Depths

Oxidation takes place at an interface that becomes gradually more and more buried and progresses further into the sample. At the same time, due mainly to the uptake of oxygen, the volume increases, raising the height of the solid surface. The situation is more complicated for bound layers, as discussed later.

For bulk oxidation, it is straightforward to calculate the way an oxide of certain thickness is distributed between increasing the surface height and eating into the substrate. For the case of bound layers, the total oxide thickness may be crudely estimated by assuming bulk behaviour and measuring height changes, albeit with respect to surrounding layers, which may also be oxidising. The evolution of the position of the oxidation front can be seen in fig 3.5.

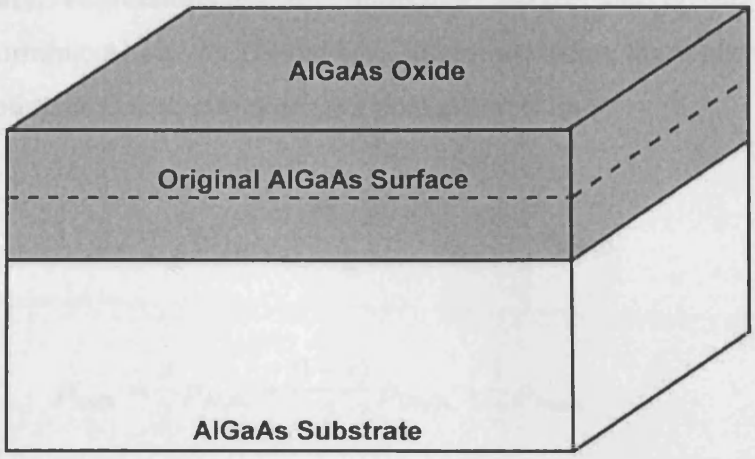

Fig 3.5 The formation of oxide on an AlGaAs substrate. The evolution of the position of the oxide/substrate interface can be clearly seen.

Consider the chemical reaction for the oxidation of $Al_xGa_{1-x}As$:

$$aAl_xGa_{1-x}As + bO_2 \rightarrow cAl_2O_3 + dGa_2O_3 + eAs_2O_3 \qquad (3.19)$$

Here, a, b, c, d and e are the molar quantities of each material involved in the oxidation reaction. This may be written in terms of the Aluminium mole fraction, x, as

$$aAl_xGa_{1-x}As + bO_2 \rightarrow \frac{x}{2}Al_2O_3 + \frac{(1-x)}{2}Ga_2O_3 + \frac{1}{2}As_2O_3 \qquad (3.20)$$

We may obtain the volume of one mole of $Al_xGa_{1-x}As$ by dividing its molecular weight, M, by its density, ρ. Assuming Vegard's law holds, these may be obtained by the following expressions:

$$M_{Al_xGa_{1-x}As} = xM_{Al} + (1-x)M_{Ga} + M_{As} \qquad (3.21)$$

$$\rho_{Al_xGa_{1-x}As} = x\rho_{AlAs} + (1-x)\rho_{GaAs} \qquad (3.22)$$

Similarly, expressions for the molecular weight and density of oxidised $Al_xGa_{1-x}As$ (forming Al_2O_3, As_2O_3 and Ga_2O_3) are, assuming that bulk values for the oxides are applicable (i.e. that they are not porous), as follows:

$$M_{oxide} = \frac{x}{2}M_{Al_2O_3} + \frac{(1-x)}{2}M_{Ga_2O_3} + \frac{1}{2}M_{As_2O_3} \tag{3.23}$$

$$\rho_{oxide} = \frac{x}{2}\rho_{Al_2O_3} + \frac{(1-x)}{2}\rho_{Ga_2O_3} + \frac{1}{2}\rho_{As_2O_3} \tag{3.24}$$

Since 1 mole of $Al_xGa_{1-x}As$ is converted to 1 mole of oxide, we may write

$$\frac{(\text{thickness of } Al_xGa_{1-x}As \text{ consumed}) \times (\text{surface area of oxidising layer})}{(\text{thickness of } Al_xGa_{1-x}As \text{ oxide}) \times (\text{surface area of oxidising layer})} = \frac{\text{molar volume of } Al_xGa_{1-x}As}{\text{molar volume of } Al_xGa_{1-x}As \text{ oxide}} \tag{3.25}$$

This assumes that the surface oxide does not spread laterally over the surface. By making this assumption it is possible to write

$$\frac{\text{thickness of } Al_xGa_{1-x}As \text{ consumed}}{\text{thickness of } Al_xGa_{1-x}As \text{ oxide}} = Z \tag{3.26}$$

where Z is the ratio of the molar volumes of $Al_xGa_{1-x}As$ and $Al_xGa_{1-x}As$ oxide.

The thickness of $Al_xGa_{1-x}As$ consumed during oxide growth is given by

$$Z \times (\text{thickness of } Al_xGa_{1-x}As \text{ oxide}) \tag{3.27}$$

The total oxide thickness, H, can therefore be related to the increase in height, h, by the following equation:

$$H = \frac{h}{(1-Z)} \tag{3.28}$$

The Z values for selected Al mole fractions are displayed in table 3.2. Thus a $Al_{0.5}Ga_{0.5}As$ layer with a height change of 10 nm measured with respect to a material that does not oxidise will have a total thickness of 33.34 nm.

***Table 3.2** Z values for selected compositions of $Al_xGa_{1-x}As$.*

Compound	Z
AlAs	0.727
$Al_{0.9}Ga_{0.1}As$	0.717
$Al_{0.8}Ga_{0.2}As$	0.710
$Al_{0.7}Ga_{0.3}As$	0.709
$Al_{0.5}Ga_{0.5}As$	0.700
$Al_{0.3}Ga_{0.7}As$	0.692
$Al_{0.1}Ga_{0.9}As$	0.684
GaAs	0.682

It should be noted that lateral oxidation studies on VCSEL devices that achieve oxidation depths of tens of microns reveal that only a small amount of oxide emerges from the edge [19]. The volume increase of the oxide must therefore relieve itself by increasing the separation of the layers, as illustrated in fig 3.6 for an oxidising AlAs layer sandwiched between two GaAs layers.

A simple calculation shows that when AlAs totally oxidises, an increase in volume of ~ 140 % occurs, assuming that no As loss occurs. Thus, a 50 nm thick layer of AlAs can be expected to increase in thickness to 70 nm on oxidation.

There are, however, several complications to the idealised assumptions made above. Firstly, it is to be expected that GaAs will oxidise at the interface. The oxide is also of a porous nature, so the above calculated increase in volume may not be correct. A further complication at higher temperatures is the fact that As is probably lost from the oxidising layer.

For this study, it is therefore expected that the increase in volume manifests itself in a combination of oxide growth at the surface and a change in layer separation.

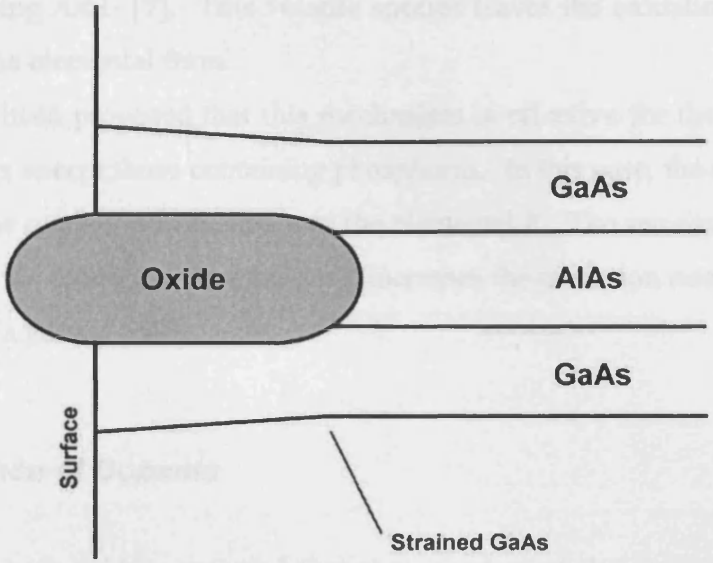

Fig 3.6 Change in layer separation caused by the growth of an oxidised region of an AlAs layer, sandwiched between two GaAs layers.

3.3.6 Presence of Moisture

Since experiments were carried out under ambient conditions, the possibility that moisture in the atmosphere will affect the oxidation processes must be considered. Indeed, it is widely reported that humid air accelerates oxide growth for GaAs, GaP, InSb and InP [20] and thus, the thickness of oxides grown by exposure to air will vary with the humidity of the room. In addition, III-V and silicon oxidation is often performed in steam.

H_2O diffuses and decomposes differently to O_2 modifying the oxidation rate. On decomposition, reactive hydroxide ions are produced via the following reaction:

$$H_2O \longrightarrow H^+ + OH^- \qquad (3.29)$$

Studies of AlGaAs based VCSEL structures at high temperature have also led to the proposition that hydrogen ions, produced as a by-product of the reaction between the group III elements and H_2O at the oxidation front, aid in the removal of

As by producing AsH_3 [7]. This volatile species leaves the oxidation front far more readily than the elemental form.

It has been proposed that this mechanism is effective for the oxidation of all III-V materials except those containing phosphorus. In this case, the ready production of $InPO_4$ at the oxidation front removes the elemental P. The removal of the group V element from the oxidation front not only increases the oxidation rate, but leads to the production of a porous oxide.

3.3.7 Presence of Dopants

It has been widely reported that the presence of dopants in silicon systems enhances oxidation [21]. In the case where *n*-type donors (e.g. phosphorus) are present, their incorporation into the oxide is known to weaken its bond structure. This acts to lower the potential barrier within the oxide, aiding the diffusion process and allowing a higher oxidant flux at the oxidation front. The presence of *p*-type donors (e.g. boron) also increases the oxidation rate, although the mechanism is less well understood. It is possible that the *p*-type donor affects the reaction process directly as a result of the incomplete bonds it forms with Si [21].

Recent research has shown that the activation energy for the oxidation of GaAs and AlGaAs are also influenced by the presence of dopant atoms within the material [10,22]. In this case, the rate of enhanced oxidation differs between the two dopant types. The upper *p*-type Bragg stack was observed to oxidise faster than the lower *n*-type stack. The explanation offered is that doping levels in the oxide strongly influence its Fermi level. This increases the number of charged defect densities that are highly dependent on the Fermi level in the oxide. Enhanced diffusion due to the elevated number of these crystal defects then occurs, thereby increasing the oxidation rate.

However, it should be noted that the effects of differences in dopant types were not observed to influence the oxidation rates of layers in any of the samples investigated in this study. It is unlikely, therefore, that the presence of dopants have a significant effect on the oxidation rate of the materials investigated during this study under the conditions at which the experiments were carried out.

3.4 Summary

The above discussion demonstrates the gaps in current understanding of the oxidation of III-V materials. It is unlikely that it is possible to gather data from the activated adsorption regime as this stage of oxidation occurs extremely rapidly. However, it may be possible to observe the other oxidation mechanisms. Each different oxidation mechanism has a different temporal dependence. These dependencies may thus be used to check for each mechanism.

With reference to equation 3.1, it can be seen that for field-aided oxidation, a $1/h \propto \log_{10}(t)$ dependence is expected. This form of oxidation is expected for bound $Al_xGa_{1-x}As$ layers for up to $x = 0.7$. For diffusion limited oxidation, a $h^2 \propto t$ dependence will be observed. This mechanism is expected to dominate for bound structures when field-aided oxidation becomes insignificant. Finally, for reaction rate limited oxidation, a $h \propto t$ dependence occurs.

3.5 References

[1] C W Wilmsen (Ed.). *Physics and Chemistry of III-V Compound Semiconductor Interfaces*. Plenum Press (1985).

[2] W Mönch. Surface Science **168**, 577 (1986)

[3] M R Brozel & G E Stillman (Eds.). *Properties of GaAs*. IEE Inspec. (1996).

[4] F Bartels, H Gröll & W Mönch. Journal of Vacuum Science and Technology, **B4**, 1100 (1986).

[5] F Cabrera & N F Mott. Reports on the Progress of Physics, **12**, 163 (1948).

[6] B E Deal & A S Grove. Journal of Applied Physics, **36**, 3770 (1965).

[7] C I H Ashby, J P Sullivan, K D Choquette, K M Geib & H Q Hou. Journal of Applied Physics, **82**, 3134 (1997).

[8] W Mönch. Springer Series in Surface Science, Vol. 26, ISBN 3-540-54423-2.

[9] R L Naone & L A Coldren. Journal of Applied Physics Letters, **82**, 2277 (1997).

[10] M Ochiai, G E Giudice, H Temkin, J W Scott & T M Cockerill. Applied Physics Letters, **68**, 1898 (1996).

[11] F Reinhardt, B Dwir, G Biasiol & E Kappon. Applied Surface Science, **104/105**, 529 (1996).

[12] D A Porter & K E Easterling. *Phase Transformations in Metals and Alloys (2^{nd} Ed.)*. Chapman and Hall (1992).

[13] K Kato, K M Geib, R G Gann, P Brusenback & C W Wilmsen. Journal of Vacuum Science and Technology, **A2**, 588 (1984).

[14] G M Yang, D H Lim, J-H Kim, K Y Lim & H J Lee. Journal of Applied Physics, **37**, 1391 (1998).

[15] K A Bertness, J-J Yeh, D J Freidman, P H Mahowald, A K Wahi, T Kendelewicz & W E Spicer. Physics Review B, **63**, 5046 (1988).

[16] H Nickel. Journal of Applied Physics, **78**, 5210 (1995).

[17] R S Burton & T E Schlesinger. Journal of Applied Physics, **76**, 5503, (1994).

[18] F A Kish, S J Caracci, N Holonyak Jr., K C Hsieh, J E Baker, S A Maranowski, A R Sugg & J M Dallasasse. Journal of Electronic Materials, **21**, 1133 (1992).

[19] T Langenfelder, St Schröder & H Grothe. Journal of Applied Physics, **82**, 3548 (1997).

[20] W E Spicer, I Lindau, P Pianetta, P W Chye & C M Garner. Thin Solid Films, **56**, 1 (1979).

[21] J Mayer & S S Lau. *Electronic Materials Science: For Integrated Circuits in Si and GaAs*. Macmillan Publishing Company (1990).

[22] F A Kish. Applied Physics Letters, **60**, 3165 (1992).

Chapter 4

Atomic Force Microscopy

4.1 Introduction

Atomic Force Microscopy (AFM) is one of many imaging techniques described generically as Scanning Probe Microscopy (SPM). Despite being a relatively recent development, SPM techniques are constantly being advanced by the demands continuous research in the field of surface science place on them.

SPM employs different surface imaging techniques than those of more traditional methods, such as electron microscopy and optical microscopy, by using a sharp tip or probe. Its ability to map the surface of samples in three dimensions offers much less ambiguity as to the true topography of sample surfaces.

The SPM era dawned with the advent of Scanning Tunnelling Microscopy (STM) [1]. STM studies surfaces by monitoring a tunnel current between a probe held over a sample and its surface. The technique allows the surface topography and electronic properties of samples to be measured. Although STM offers excellent resolution (down to atomic scales), it can only be used to probe electrically conductive samples.

In response to this severe limitation, Binning, Quate and Gerber developed the first AFM in 1986 [2]. Here, the surface is studied by means of the force interactions between a probe and the sample surface. Thus, the range of materials that SPM techniques can be used to study has been greatly expanded. As well as conductive materials, AFM can be used to characterise the surfaces of insulators [3], and can even operate under liquids [4], making it particularly effective for studies of biological specimens [5]. The principles behind AFM have also lead to the measurement of surface properties such as friction, magnetic forces, elasticity and stiffness [6, 7]. It is difficult to say whether or not AFM is superior to its precursor STM. The use of a tunnelling current allows STM to achieve a lateral resolution of 1 - 2 Å [8], marginally better than that capable by AFM, whose resolution is restricted

by the 'contact' area of the tip. However, STM's restriction to conductive samples makes it far less versatile.

This chapter explores various aspects of AFM technology. Section 4.2 gives a brief introduction to the principles of AFM operation, and section 4.3 goes on to discuss the different modes of operation in detail. Section 4.4 describes the types of cantilevered tips that are in general use today, while section 4.5 gives a brief overview of scanner design and some of the problems encountered in their operation. Feedback controls and scanning parameters are discussed in section 4.6. Finally, section 4.7 deals with the Digital Instruments NanoScope™ IIIa MultiMode™ SPM. All the AFM work in this research was carried out using this model, and it should be noted that all discussions in this chapter refer to this instrument.

4.2 The Fundamentals of Atomic Force Microscopy

An AFM probe consists of an atomically sharp tip, typically a few microns long and with a radius of curvature ~ 5 nm [9]. The tip is mounted on the free end of a flexible cantilever, 100 to 200 μm long [9], that extends from a rigid substrate.

Samples are mounted on a piezoelectric *xyz* scanner that allows precise motion in any direction. The cantilever bends in response to any forces that arise when the tip interacts with the sample surface. In the simplest mode of operation, these deflections are monitored by a feedback system that maintains a constant tip-sample interaction by adjusting the voltage that is applied to the piezoelectric scanner. As the tip is systematically scanned, or *rastered*, across the surface, a complete three-dimensional image of the sample surface is built up by monitoring this adjustment voltage [10]. This image is then displayed in false-colour on a computer screen. The experimental arrangement is shown in fig 4.1.

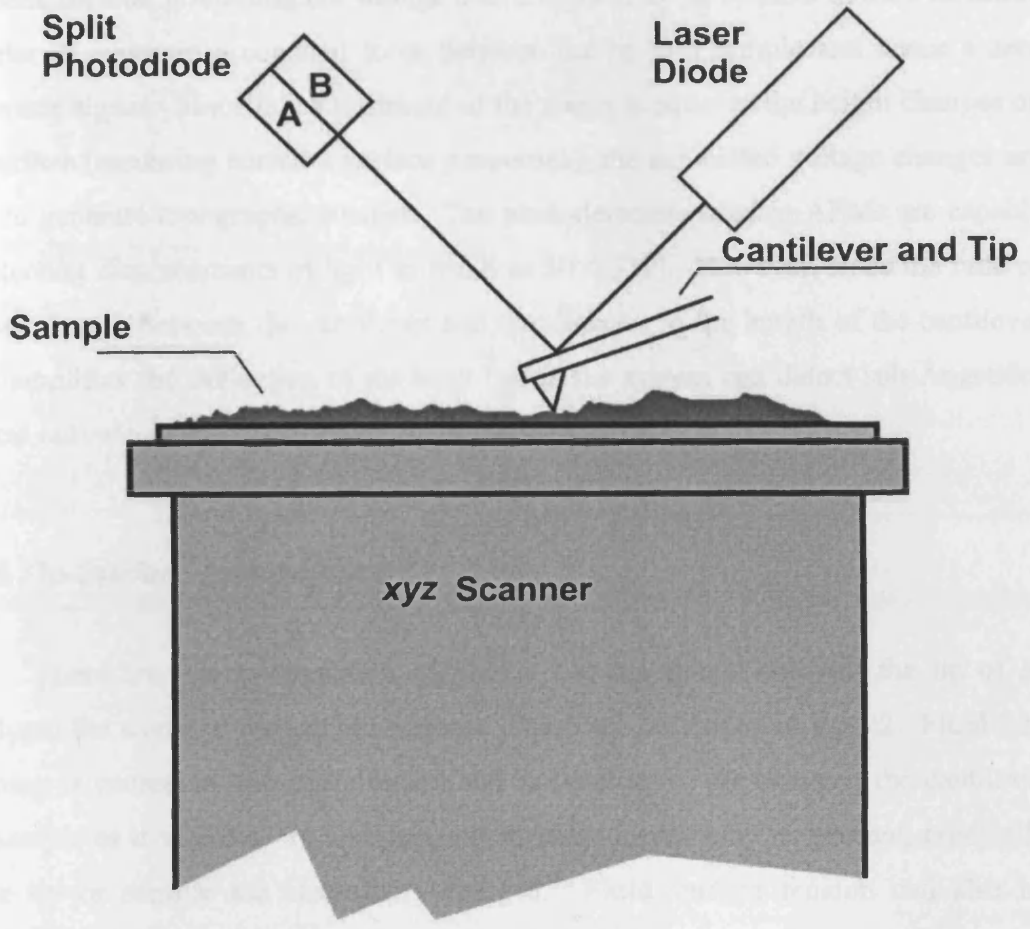

Fig 4.1 Schematic diagram of the AFM set-up.

4.2.1 Cantilever Deflection Detection

There are several techniques by which the deflection of a cantilever may be determined. The Digital Instruments NanoScope™ IIIa MultiMode™ SPM uses the *optical lever method* [11]. A laser beam is reflected off the back of the cantilever and is incident on the interface of a split photodiode array. The array is divided into two equal segments, A and B, as shown in fig 4.1. The difference signal (A-B), between the two segments of the photodiode for no deflection is set to zero. Any subsequent deflections caused by height changes over the surface of the sample result in non-zero values of (A-B).

In the simplest mode of operation, the difference signal is used directly as a feedback control, governing the voltage that is applied to the scanner in the *z* direction in order to maintain a constant force between the tip and sample and hence a zero difference signal. Since this adjustment of the *z*-axis is equal to the height changes on the surface (assuming constant surface properties), the associated voltage changes are used to generate topographic images. The photodetectors used in AFMs are capable of detecting displacements of light as small as 10 Å [12]. However, since the ratio of the path length between the cantilever and the detector to the length of the cantilever itself amplifies the deflection of the laser beam, the system can detect sub-Ångström vertical movement of the cantilever tip at the sample surface.

4.2.2 Tip-Surface Interactions

There are many interactive processes that can occur between the tip of an AFM and the atoms of the sample surface. These are illustrated in fig 4.2. Fluid film damping is caused by the compression and rarefaction of air between the cantilever and sample as it vibrates. In addition, electrostatic forces may be present, especially if the tip or sample are electrically charged. Fluid surface tension can also be important and is due to the capillary force that results from a microscopic layer of adsorbed water layer on the sample surface [13]. However, fluid surface tension only acts on the tip as it retracts from the surface, trying to pull it back down. In the absence of moisture, van der Waals forces and Coulomb repulsion dominate [14].

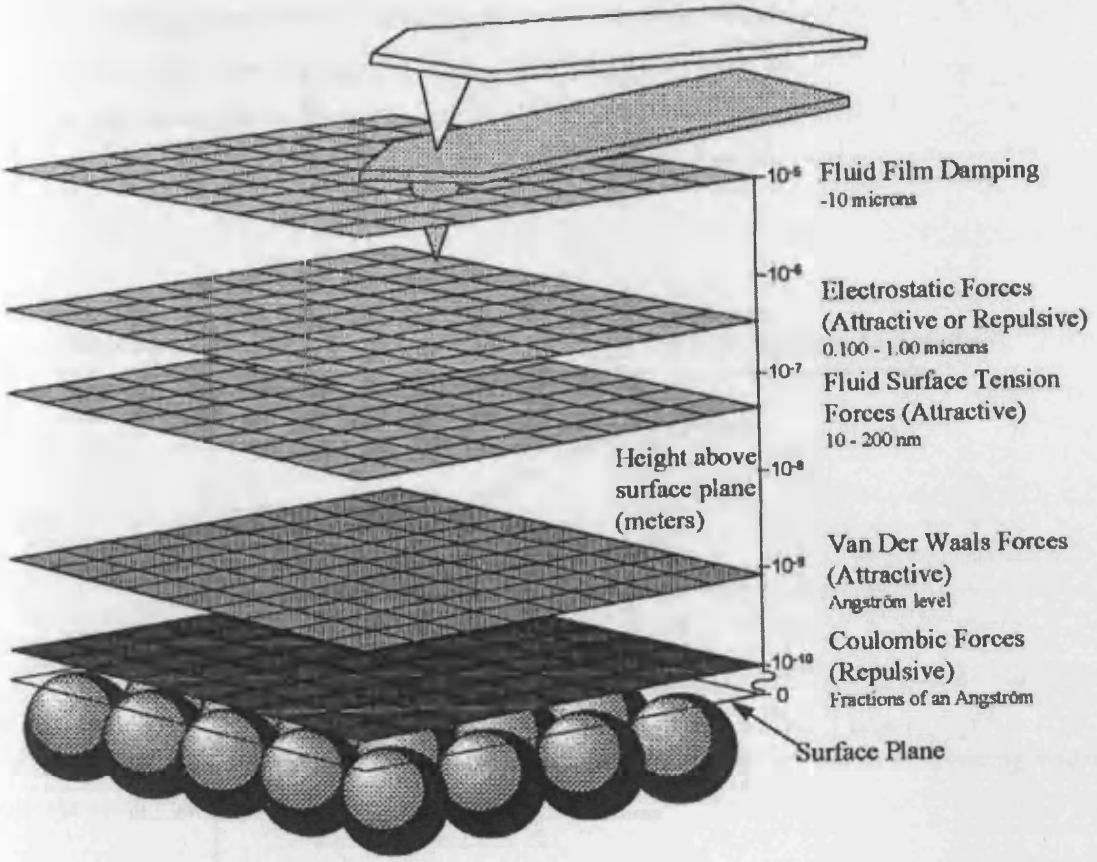

Fig 4.2 *The interaction forces experienced by a tip as it approaches a sample surface [9].*

The van der Waals attractive force results from a combination of several effects. The main attractive contribution is from the dispersion force that results from the induced dipole interaction of electrons around the nucleus.

The second repulsive interaction, Coulomb repulsion, results from the overlapping of the electron clouds of atoms as the tip and sample are brought close together. If the two interactions are combined, the well known *Lennard-Jones Potential* curve is obtained, as shown in fig 4.3.

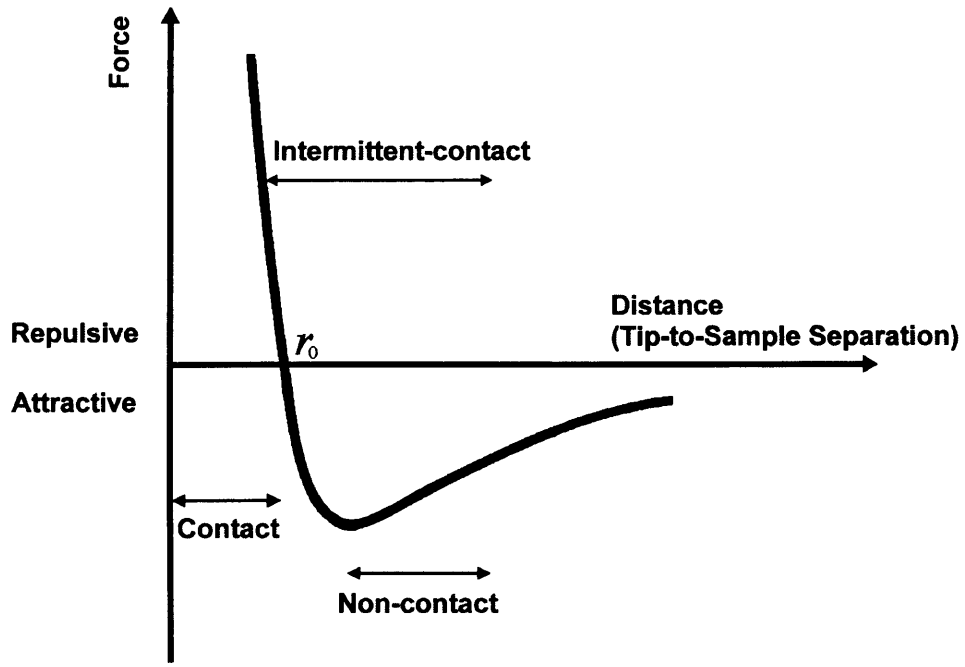

Fig 4.3 The Atomic force-distance curve, showing the regimes for the various AFM operating modes. r_0 is the equilibrium separation.

Fig 4.3 distinguishes three distinct interaction regimes (corresponding to the three different modes under which AFM can operate): *contact, non-contact* and *intermittent-contact.* These are dealt with in section 4.3.

4.3 Modes of Operation

4.3.1 Contact Mode Atomic Force Microscopy

In contact mode, topographic images are built up by placing the tip in physical contact with the sample and dragging it across its' surface [15]. As illustrated in fig 4.3, the force curve is very steep in the contact regime, indicating that the repulsive force strongly resists any attempts to push the tip and sample close together. Cantilevers must bend in response to surface. For this reason, contact mode cantilevers are designed with spring constants lower than the effective spring constant holding the atoms together.

There are two difficulties encountered with contact mode imaging. Firstly, the contact force exerted by the cantilever itself can be quite destructive, causing damage to the tip and sample. It is therefore important to utilise cantilevers with optimised spring constants, as outlined above.

Secondly, the layer of adsorbed water vapour may cover the surface may be up to several nanometers thick. The capillary action of this layer can distort measurement data and also damage both the tip and sample. However, because the distance between the tip and the sample is virtually incompressible, the capillary force should remain constant provided the tip is in permanent contact with the surface and the water layer is homogeneous. The variable force in contact AFM is therefore the force exerted by the cantilever.

4.3.2 Non-Contact Mode Atomic Force Microscopy

Non-contact mode was developed in response to the inherent problems of sample damage associated with contact mode AFM. Here, sample topography is imaged by holding the tip a small distance (in the region of 10 to 100 Å) above the surface and detecting the attractive van der Waals forces between the tip and sample [15].

Consequently, the variations in force experienced between the tip and the sample in non-contact AFM are several orders of magnitude lower than that of contact mode, resulting in much smaller cantilever deflections and hence poorer resolution. This problem is further enhanced by the fact that cantilevers used for non-contact AFM must be stiffer than those used for contact AFM. This is because soft cantilevers can be pulled into contact with the sample surface by the attractive Van der Waals force. Thus, the non-contact signal is small and more difficult to measure. Sensitive detection schemes must therefore be employed.

During operation, the cantilever is oscillated near its resonant frequency (typically ~ 100 to 400 kHz) with an amplitude of a few nanometers [9], and is kept vibrating at this frequency with the aid of a feedback loop. Changes in resonant frequency depend on the van der Waals force encountered, and are thus sensitive to sample topography. By keeping the resonant frequency at a fixed value as the surface is scanned, the system maintains the same tip-to-sample distance. The motion of the

scanner in response to these frequency changes can therefore be used to map the surface topography.

The advantage of non-contact mode is immediately apparent. There is no contact between the tip and the sample surface. Hence the tip and sample do not suffer the degradation effects that occur during the continuous use of contact AFM. The one disadvantage associated with non-contact imaging, however, is that the van der Waals forces that the tip detects extend only a short distance. If there is an adsorbed water layer present, non-contact mode will not image the true surface topography. Whereas as tip operating in contact mode will penetrate the layer and image the underlying surface, a tip operating in non-contact mode will image the surface of the water layer. This effect is illustrated in fig 4.4. The technique is used exclusively under vacuum conditions for the above reasons.

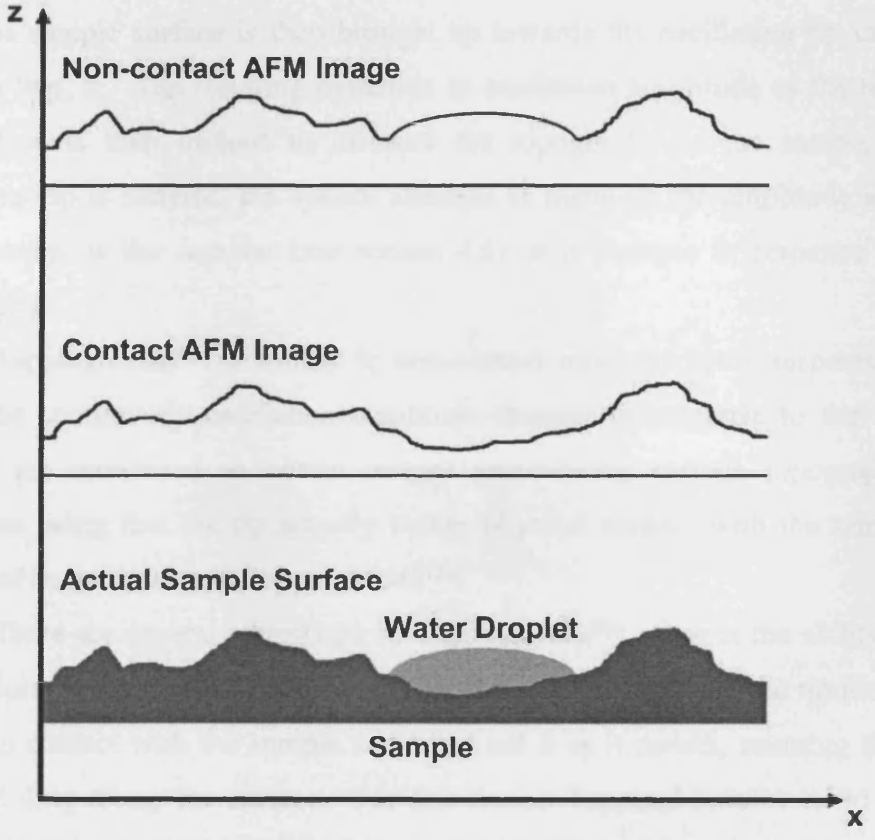

Fig 4.4 A comparison of the line scans obtained by imaging a surface using contact and non-contact mode where water contamination present.

Non-contact imaging is therefore not the ideal solution to the problems associated with contact imaging. Indeed, it is seldom used nowadays. It does however provide the link between contact mode and intermittent-contact or *TappingMode*™ AFM.

4.3.3 TappingMode™ Atomic Force Microscopy

TappingMode™ AFM was first developed by Digital Instruments, and in other operating systems is known as intermittent-contact mode. It has become an important AFM technique since it largely overcomes the limitations of both contact and non-contact AFM [9].

During operation, the cantilever is firstly made to oscillate at or near its resonant frequency (~ 300 kHz), producing a free air amplitude greater than 20 nm [9]. The sample surface is then brought up towards the oscillating tip until the tip starts to 'tap' it. The resulting reduction in oscillation amplitude as the tip contacts the surface is then utilised to measure the topography of the sample. As the oscillating tip is rastered, the system attempts to maintain the amplitude at a pre-set value known as the *setpoint* (see section 4.6) as it changes in response to surface topography.

TappingMode™ is similar to non-contact mode in many respects. In both cases, the cantilever's oscillation amplitude changes in response to the tip-sample spacing are monitored to obtain images representing surface topography. The difference being that the tip actually makes physical contact with the sample at the bottom of its oscillation in TappingMode™.

There are several advantages to TappingMode™. One is the ability to image soft samples without damaging them. This is due to the fact that the tip is alternately placed in contact with the sample and lifted off it as it rasters, ensuring that the tip does not drag along the surface. For this reason TappingMode™. AFM is widely used for imaging biological and polymer samples [9]. The amplitude of oscillation of the tip is sufficient enough to overcome any adhesion forces. Added to this, there are no shear forces, and the mode can be operated under liquids. It also has the advantage of being easy to use and requires little sample preparation. However, the lateral

resolution of TappingMode™ is debatable. It is certainly less than that of contact mode AFM, and is probably about 5 nm [12].

4.3.4 Phase Imaging

Phase imaging, or *phase detection microscopy* is a technique that may be used while an AFM is operating in TappingMode™ [9]. The AFM equipment allows for the monitoring of the phase lag between the signal that drives the cantilever to oscillate and the cantilever oscillation output signal. Any changes in phase lag reflect differences in the mechanical properties of the sample surface. This provides complimentary information to the topographic image, revealing variations in viscoelasticity, adhesion and friction of the sample surface.

Phase data is obtained from the difference between the driven and the actual oscillations of the cantilever. If the tip is driven with a force $F_0 cos(\omega_t - \phi_d)$, the actual response of the tip will be $A_0 cos(\omega_t - \phi_r)$. The phase offset caused by interaction with the surface is then given by $(\phi_d - \phi_r)$. The experimental configuration is shown in fig 4.5.

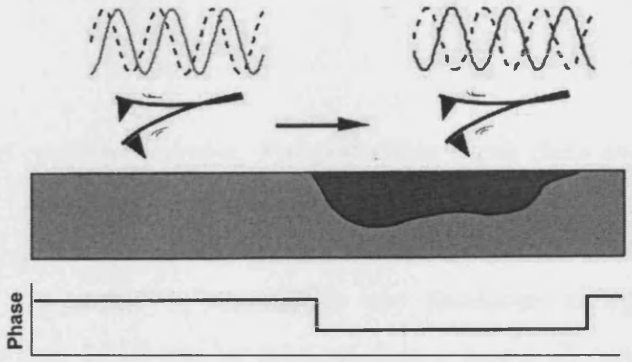

Fig 4.5 The origin of phase imaging: the cantilever driving oscillation (solid wave) is compared to that of the actual response of the tip (dashed wave). Variations in sample properties cause a phase difference between the two [9].

The system's feedback loop operates in the usual manner, using changes in the cantilever's vibration amplitude to measure sample topography. The phase lag is

monitored while the topographic image is being taken, so that images of topography and material properties can be collected simultaneously.

4.4 Cantilevered Probes

4.4.1 Design and Types of Cantilevered Probes

There are primarily two designs of cantilevered probes available for use with AFM. The first are for use in TappingMode™ and are fabricated from silicon. The second type are the more flexible contact mode cantilevers, fabricated from silicon nitride. The two types are compared in fig 4.6.

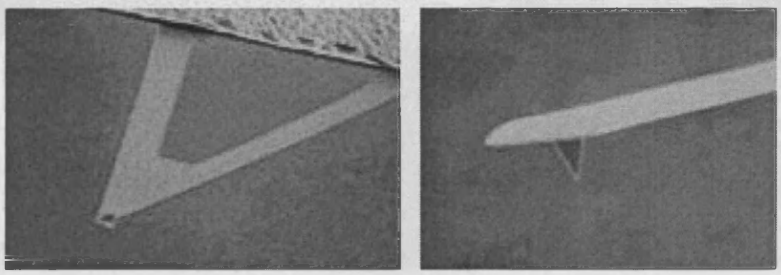

Fig 4.6 *Two types of cantilevered probes: V-shaped silicon nitride (left), and crystal silicon (right) [9].*

Integrated tip-cantilever assemblies are produced using photolithographic techniques. More than 1,000 can be fabricated on a single silicon wafer, each with a typical lifetime of a couple of days [9]. Tips are typically a few microns long, with cantilever lengths of between 100 and 200 μm [6]. The theoretical tip shape for a silicon TappingMode™ cantilever is illustrated in fig 4.7.

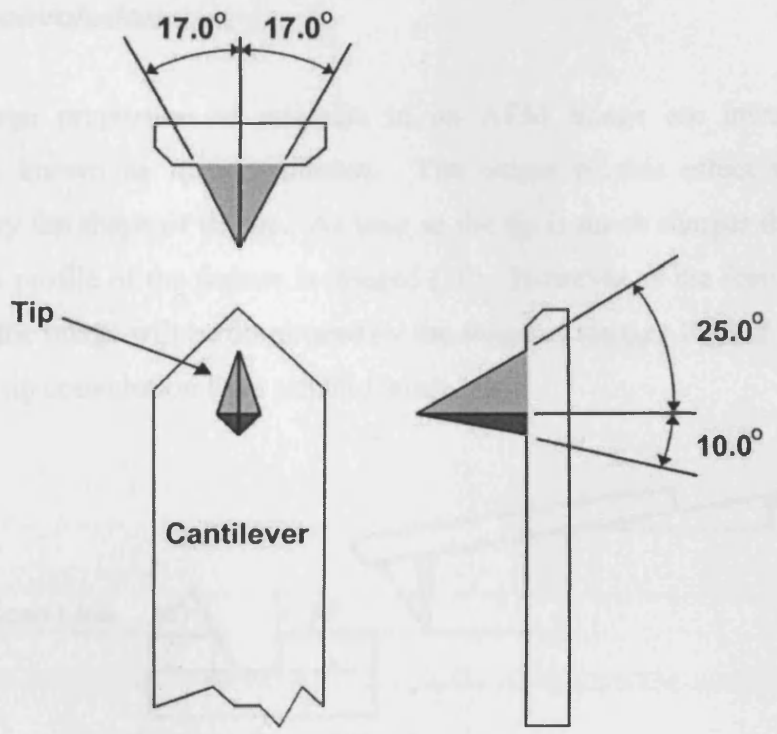

Fig 4.7 Theoretical tip shape for a silicon cantilever [9].

Ultimately, both the lateral resolution and the force applied to the sample are determined by the properties of the cantilever and its tip. For successful AFM imaging, a sharp tip and a cantilever with an optimised spring constant is required. The spring constant of the cantilever depends on its shape, its dimensions, and the material from which it is fabricated. Generally, thicker and shorter cantilevers are stiffer and have higher resonant frequencies. The spring constants of commercially available cantilevers range over four orders of magnitude, from thousands of a Newton per meter to tens of Newtons per meter [12]. Resonant frequencies range from a few kilohertz to hundreds of kilohertz [12].

The sharpest commercially available tips can have radii as small as 5 nm, although during use may be worn, resulting in typical values of ~ 20 nm [12]. However, the actual tip-surface interaction area is somewhat smaller, enabling atomic resolution under vacuum conditions.

Image size must also be considered as this too determines the resolution for a particular image. The lateral resolution of AFM images often depends not on the tip radius, but the image's step size.

4.4.2 Tip Convolution

A large proportion of artefacts in an AFM image are introduced by a phenomenon known as *tip convolution.* The extent of this effect is essentially determined by the shape of the tip. As long as the tip is much sharper than a feature, the true edge profile of the feature is imaged [16]. However, if the feature is sharper than the tip, the image will be dominated by the shape of the tip. Fig 4.8 demonstrates the origin of tip convolution for a trench feature.

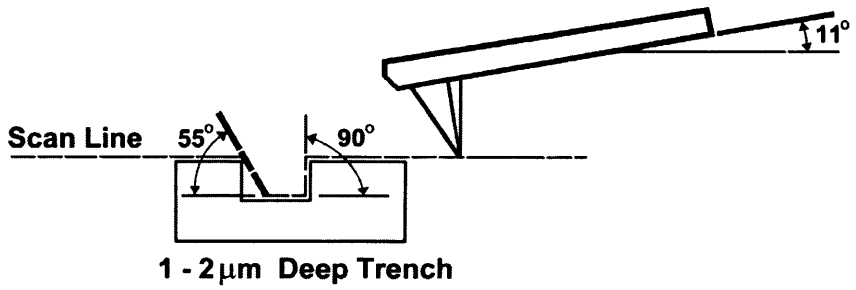

Fig 4.8 An illustration of tip convolution: The scan line produced using the theoretical probe tip shape shown in fig 2.6 on a 1-2 μm deep vertically walled trench. Any wall angle on the left wall that is greater than 55° will be shown as 55° in the image. Note also that the cantilever is not parallel with the sample surface [9].

Tip convolution can therefore be a major problem when imaging samples with steep-sided features. Provided the tip touches the bottom between features, the height of the feature is reproduced accurately, leading to fairly accurate height measurements and roughness statistics. The measured lateral dimensions however, give only a maximum value.

Of relevance to this study is that the fact that differences in oxidation rates for the different material layers of semiconductor superlattice structures cause trench like features whose width can be similar to that of the thickness of the slowest oxidising layer. If the trench width is similar to the radius of curvature of the tip, then it may not penetrate to the bottom of the trench. In this case an erroneous measurement in oxide height will result. Tip convolution should be suspected particularly if the bottoms of trenches appear rounded. A more detailed analysis of the tip convolution effects expected to be encountered during this study appears in section 5.3.4.

4.5 Scanner Characteristics

4.5.1 Design Considerations

Piezoelectric scanners are designed to move in x, y and z directions by expanding and contracting in appropriate directions under the application of an alternating voltage. The piezoelectric material is a polycrystalline ceramic, each of the crystals having individual, randomly oriented dipole moments. If a direct voltage is applied to the heated material (~ 200 °C), the dipoles become aligned [6]. Quick cooling of the material preserves the alignment, allowing the scanner to change in response to applied voltages.

The AFM scanner utilises a simple tube design, as illustrated in fig 4.9. The maximum scan size that can be achieved by a particular piezoelectric scanner is dependent on a number of factors, including the length and diameter of the scanner tube, its wall thickness and the strain coefficients of the particular piezoelectric material from which it is fabricated [12]. AFM scanners are available that can scan laterally from tens of nanometers to 200 μm and vertically from the sub-Ångström range up to 8 μm [12].

The scanner is used as an extremely fine positioning stage to move the sample under the probe. The AFM electronics drive the scanner in a raster pattern, shown in fig 4.10. The scanner moves across the first line of the scan, and back. It then steps up in the perpendicular direction to a second scan line, moves across it and back, then to the third line, and so forth. AFM data are only usually collected in the direction parallel to the motion of the tip, known as the *fast-scan direction.* The perpendicular direction, in which the scanner steps from line to line, is called the *slow-scan direction.*

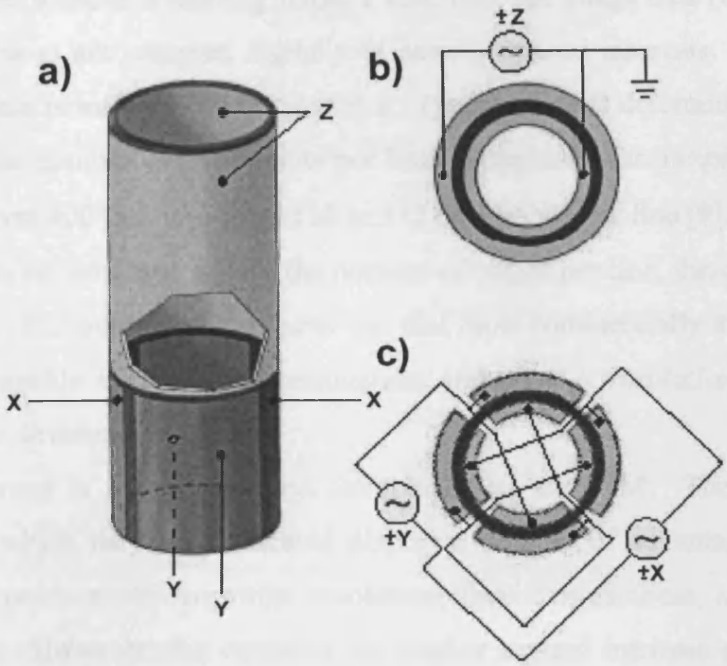

Fig 4.9 (a) A typical scanner piezo tube and its xyz electrical configurations (b) internal z electrical configuration (c) internal x and y electrical configuration. AC signals applied to conductive areas of the tube create piezo movement along the three major axes [9].

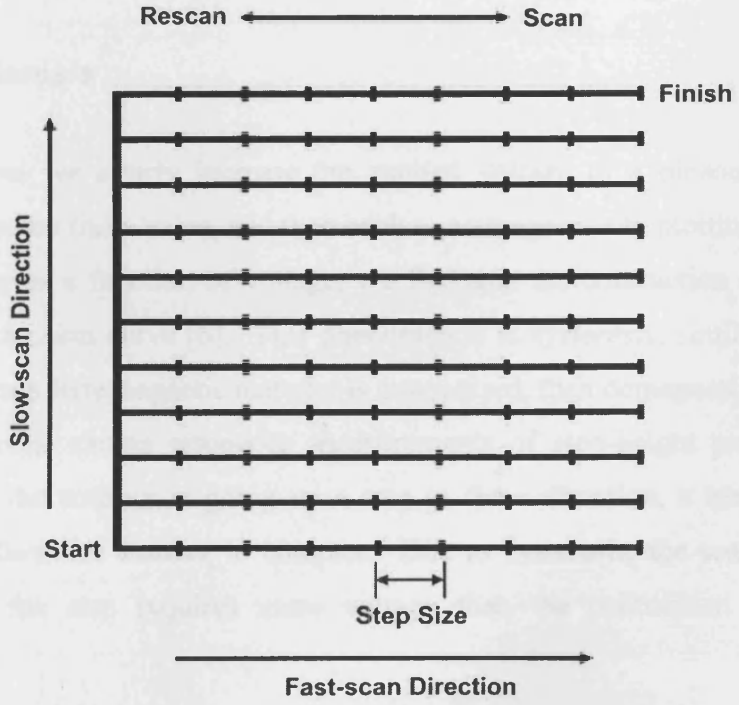

Fig 4.10 Scanner motion during data acquisition.

While the scanner is moving across a scan line, the image data (e.g. the height of the scanner in z) are sampled digitally at equally spaced intervals. The spacing between each data point is called the *step size*. The step size is determined by the full scan size and the number of data points per line. Scan sizes can range from tens of Ångströms to over 100 μm, and from 128 to 512 data points per line [9]. The number of lines in a data set generally equals the number of points per line, thereby producing a square image. It should be noted, however, that most commercially available AFM equipment is capable of producing rectangular images. A resolution of 0.1 Å is possible in the z direction.

The scanner is therefore a vital component of an AFM. The piezoelectric ceramics from which they are fabricated display a number of advantages, including their ability to produce sub-Ångström resolution, their compactness, and their high-speed response. However, the ceramics do display several intrinsic properties that warrant attention. Specifically, they inhibit the ability to very accurately determine the dimensions of features on a sample. These properties are outlined below.

4.5.2 Scanner Nonlinearities

4.5.2.1 Hysteresis

Suppose we slowly increase the applied voltage to a piezoelectric scanner from zero to some finite value, and then back to zero again. On plotting the extension of the scanner as a function of voltage, we find that the contraction curve does not retrace the extension curve [6]. This phenomenon is *hysteresis*, similar to the effect observed when a ferromagnetic material is magnetised, then demagnetised.

Hysteresis causes erroneous measurements of step-height profiles in the z direction. If the scanner is going up a step in the z direction, a certain voltage is required to allow the scanner to contract. Due to hysteresis, the scanner extension going down the step requires more voltage than the contraction for the same displacement.

4.5.2.2 Creep

Abrupt changes in voltage give rise to what is known as *creep* [12]. This is a result of the inability of the piezoelectric material to change dimensions instantaneously. In fact, dimension changes occur in a two-step process. The first stage, where up to 99 % of the change occurs, takes place in less than a millisecond. The second stage takes place over a considerably longer time scale, perhaps tens of seconds, whereby the scanner gradually assumes the required dimensions.

Creep generally has the effect of 'rounding-off' the corners of the sharp features of a sample. This is obviously exacerbated by large steep sided features that require large voltage changes to scan. Creep can also manifest itself in the lateral dimensions when a large abrupt change in scan dimensions occurs.

4.5.2.3 Ageing

Recall from section 4.5.1 that piezoelectric materials are polycrystalline ceramics, with each of the tiny crystals having its own dipole moment. The amount of deflection achieved by the scanner depends on how many of these dipole moments are aligned. However, not all of the crystals are aligned when the scanner is new. Repeated application of voltage in the same direction results in more and more of the dipoles aligning themselves. Thus, with regular use, the deflection achieved for a given voltage increases slowly with time [12].

Conversely, if the scanner is little used, the tendency of the dipoles to gradually become randomly orientated results in fewer dipoles to contribute to the deflection of the scanner. It is therefore essential to regularly calibrate the AFM scanner to ensure that the measured dimensions of features are correct.

4.5.2.4 Cross Coupling

In order to produce x-y motion, one side of the scanner must be contracted while the other expands. Thus, a spurious movement in the z direction arises, referred to as *cross coupling* [12]. In order to maintain the AFM tip at a fixed distance from the sample surface, the piezoelectric tube must therefore scan in an arc. This can

result in the spherical distortion of images obtained from flat samples. The effect becomes particularly conspicuous for larger scans, where the length of the arc traced by the scanner is relatively large.

4.5.2.5 Nonlinearities in Combination

To have any one of the above scanner nonlinearities operating in isolation is virtually impossible. Under normal operating conditions, all of these effects will be present and will effect images by varying degrees. Fig 4.11 shows the sum of the effects of hysteresis, creep and cross coupling on the scan line of a single step. Added to this, the aspect ratio of the tip must be taken into consideration, as it may also effect the shape of any features imaged.

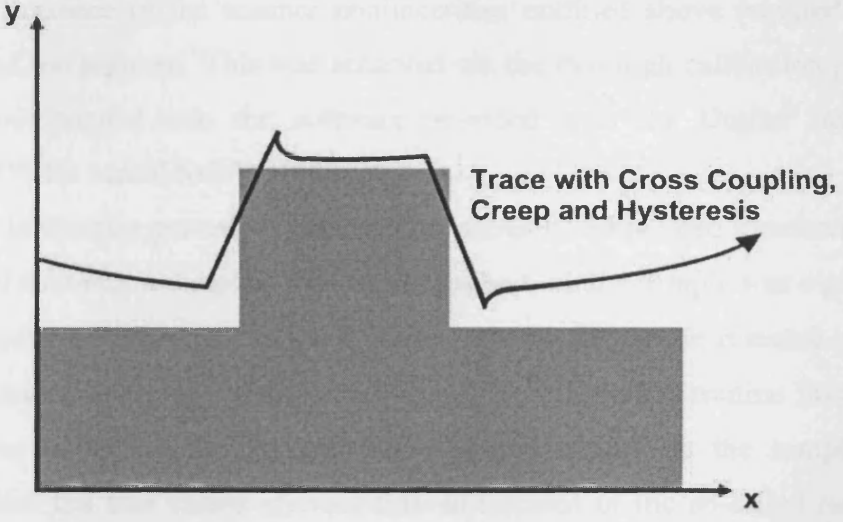

Fig 4.11 The scan line produced over a step by a scanner displaying the combined effects of hysteresis, creep and cross coupling.

Traditionally, software solutions have been employed to correct the scanner nonlinearities. These do not always prove sufficient however, and more recently, hardware solutions have become available that eliminate nonlinearities instead of correcting them. The best systems on the market today utilise a combination of hardware and software corrections, drawing on the strengths of each.

Despite the fact these methods are excellent at eliminating the above problems, it is good practice to routinely check for any image artefacts, which may be caused not only by the scanner but the tip. The following protocol was therefore used when taking images:

1. Repeat the scan to ensure it looks the same.
2. Change the sample orientation by 180 ° and take a new image.
3. Change the scan size and take an image to see if the features scale properly.
4. Change the scan speed and take another image (particularly if any suspicious periodic features appear in the image).

4.5.3 Scanner Calibration

The presence of the scanner nonlinearities outlined above required periodic calibration of the scanner. This was achieved via the thorough calibration procedure that was incorporated into the software provided with the Digital Instruments NanoScope™ IIIa MultiMode™ SPM.

The calibration procedure consisted of series of automated measurements on two standard calibration samples. For lateral calibration, the sample was a grid with a guaranteed pitch of 2 μm, so chosen because this was the same distance regime at which the device metrology was carried out. The calibration routine involved the determination of values for various distances and angles on the sample. Any deviation from the true values allowed fine adjustment of the so-called *calibration coefficients* for the scanner. In order to calibrate the scanner in the z direction, atomic step heights were measured on a standard graphite calibration sample [17].

After the calibration procedure had been completed, the samples were re-measured to ensure successful calibration had been achieved. In this case, an accuracy in measurement of 2 - 3 % could be expected from the equipment both laterally and vertically.

4.6 Scanning Parameters and Feedback Controls

The optimisation of the various scanning parameters and feedback controls is vital for obtaining good quality AFM images [10]. For example, high frequency noise may be generated in images due to oscillations in the system caused by setting the feedback gains too high.

On the other hand, if the feedback gains are too low, the tip has difficulty tracking the surface. In extreme cases, the image looses detail, appearing smooth or 'fuzzy'. A less obvious effect is 'ghosting'. On sharp slopes, an overshoot can appear in the image as the tip travels up the slope, and an undershoot can appear as the tip travels down the slope. This feedback artefact commonly appears on steep features, represented as bright ridges on the uphill side and dark shadows on the downhill side.

There are several parameters and controls that require careful consideration when using TappingMode™. These are outlined below.

Setpoint is the value of the RMS of the cantilever vibration amplitude that the feedback loop maintains. When scanning, the cantilever oscillation should be maintained at this setpoint, usually 50 - 80 % of free air amplitude. Any deflection from the setpoint causes a voltage to be applied to the piezoelectric scanner returning it to its value.

Drive Amplitude is the amplitude of the voltage at which the cantilever is driven.

The **Proportional** and **Integral Gains** determine how sensitive the feedback loop is to variations in the tip's amplitude of oscillation. Proportional Gain determines the restoration voltage applied to the cantilever for a given error and is proportional to the error in the probe position. Integral gain offers further improvement in the feedback system by averaging the error over several position readings to give a lower error by addressing error over a longer period of time. This tends to smooth out short-term fluctuations found with proportional gain, keeping the probe closer to the setpoint. For normal scanning, integral gain alone provides satisfactory control. Increasing

these gains often produces better images but only up to a point, above which high frequency noise is observed.

Scan Rate controls the rate at which the cantilever scans across the sample area. A slower scan rate generally leads to better images, but not always.

The **Number of samples** determines the number of pixels used to create the image. Increasing this value leads to a better quality image, but the time taken to obtain the image is longer.

4.7 The Digital Instruments NanoScope™ IIIa MultiMode™ Scanning Probe Microscope

4.7.1 Equipment Characteristics

The Department of Physics and Astronomy at The University of Wales, Cardiff took delivery of a Digital Instruments NanoScope™ IIIa MultiMode™ SPM in September 1997. Since that time it has proved itself an invaluable tool in many research projects throughout the Department. It is known as the MultiMode™ SPM since it offers a wide range of STM and AFM techniques. The SPM hardware is shown in fig 4.12.

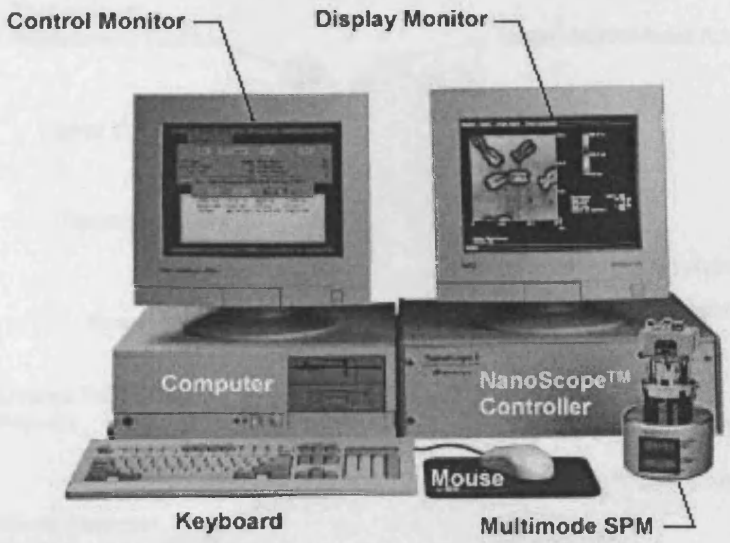

Fig 4.12 MultiMode™ SPM system hardware [9].

The NanoScope™ Controller unit is the interface between the MultiMode™ SPM and the control computer. Two monitors allow the user to control scanning via the control monitor while viewing images as they are captured on the display monitor.

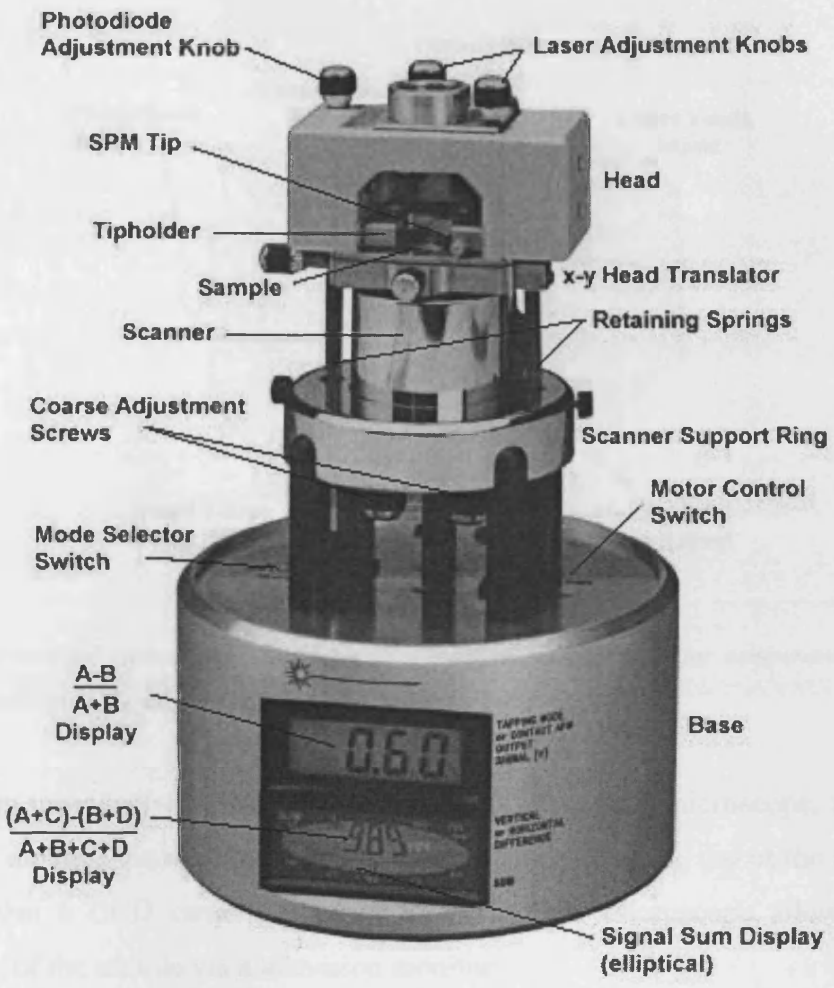

Fig 4.13 The MultiMode™ SPM [9].

The MultiMode™ SPM unit itself, shown in fig 4.13, consists of a base containing control electronics and calibration displays that give information about the laser, photodiode, and cantilever deflection. Also mounted in the base is a coarse control system that is used to move the sample over large (mm) distances before the piezoelectric scanner is used. Mounted above the base is the piezoelectric *xyz* scanner. At Cardiff, two scanners are available: The 'J' scanner (125 x 125 x 5 μm) and the smaller 'E' scanner (10 x 10 x 2.5 μm) [9]. Above the scanner is the SPM head that contains the laser, photodiode and tip holder, a schematic of which is found in fig 4.14.

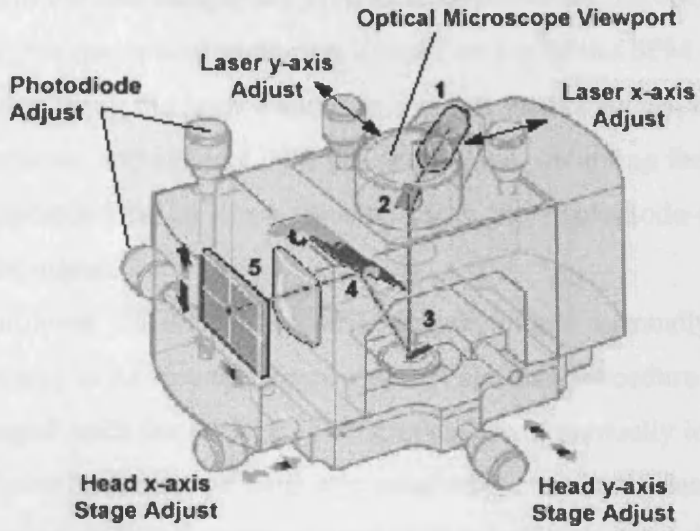

Fig 4.14 An internal view of the MultiMode SPM head, revealing its major components: laser (1); mirror (2); cantilever (3); tilt mirror (4); photodetector (5) [9].

Also provided with the NanoScope™ was an optical microscope, that allows the user to monitor the sample through a viewport located at the top of the SPM head. There is also a CCD camera attached to the optical microscope allowing easier monitoring of the sample via a television monitor.

Throughout this research, the NanoScope™ was placed on an air table, which acted as a vibration isolation system. This, in turn, is placed in a sound-absorbing room in the basement of the building. The temperature and the relative humidity of the air surrounding the equipment were constantly monitored and were found to remain in the range 17 – 21 °C and 45 - 65 % respectively.

4.7.2 Equipment Operation

All AFM imaging for this research was carried out using TappingMode™. The following section gives a brief account of its use.

Samples were mounted on a specially constructed puck that allowed cross-sectional studies to be carried out on the wafers with ease. The design of the puck is outlined in more detail in section 5.2.4. The puck, fabricated from steel, is held to the top of the scanner by a permanent magnet. Once the puck is in place, the tip holder is fitted above the sample, where it is held in place by a two-pronged grip.

The cantilever and sample are viewed from above by the optical microscope. This is achieved via the optical view-port located on top of the SPM head. The laser beam is positioned using the laser x and y axis-adjust screws so that it illuminates the back of the cantilever and reflects onto the photodiode, obtaining the maximum sum readout. The photodiodes are then adjusted using the photodiode-adjust screws to give a zero (A-B) signal.

The cantilever oscillation is then tuned (either manually or using the 'autotune' software) to its resonant frequency. When this procedure is complete, the tip can be 'engaged' with the surface. The first step is to manually lower the tip with the aid of the optical microscope until it is relatively close to the surface. The tip is then further lowered under computer control until it begins to 'tap' the surface and is able to scan the sample. Once the feedback controls have been optimised for the surface in question, measurements can be taken.

4.8 Atomic Force Microscopy Studies of Semiconductors

Almost from its conception, AFM has been used to study semiconductor materials. Indeed, the limitations of existing SPM techniques in this area of study contributed to the motivation that drove the development of AFM. Any literature review will reveal an enormous number of semiconductor studies carried out using AFM. In order to put the work of this thesis in context, chapter 4 concludes with a brief summary of current and past research in the area.

Initial work on semiconductors using AFM concentrated on epitaxially grown surfaces [18], with results produced being quantitavely similar to those obtained using other SPM techniques. Since it was established that the growth of oxide on the epitaxial surface does not impede the ability to visualise its morphology, even down to monolayer scales [19], AFM in air has become the technique of choice for a vast number of these studies.

Since these initial studies, AFM has been used as a tool to carry out a wide variety of studies on semiconductor materials, especially in conjunction with other SPM techniques in a complimentary manner. However, it is only in recent years that advances in AFM technology have really brought it to the forefront of such research.

New and exciting AFM-based techniques such as scanning capacitance and electric force microscopy have been developed that allow semiconductor materials to be probed even further. For example, it is now possible to carry out routine dopant profiling of devices [20] providing device manufacturers with an extremely powerful characterisation technique.

Dopant profiling requires the cross-sectional analysis of devices. Recently, attention has turned to the possibility of using AFM as a metrology tool in order to obtain both thickness and composition measurements of the individual epitaxial layers that constitute a semiconductor device. As discussed in previous chapters, this is already possible with existing SPM techniques, but complicated and lengthy sample preparation times make them difficult to use as routine metrology tools in an industrial environment.

In order to make these measurements with AFM, two possible layer contrast mechanisms have been considered: that of differential oxidation [21] and selective etching [22]. In both cases, a height difference is produced between the individual layers of the device structure, and the measured topography translated into layer thickness and composition data.

In this thesis, the differential oxide contrast mechanism is heavily explored after initial selective etching studies produced unsatisfactory results. It is believed that the work presented in this thesis is the first to attempt to thoroughly quantify this technique and establish its potential for the routine metrology of III-V device structures.

4.9 References

[1] New Scientist. 23rd October 1986.

[2] G Binning, C F Quate & Ch Gerber. Physics Review Letters, **56**, 930 (1986).

[3] T R Albrecht & C F Quate. Journal of Applied Physics, **62**, 2599 (1987).

[4] O Marti, B Drake & P K Hansma. Applied Physics Letters, **51**, 484 (1987).

[5] P K Hansma, V B Elings & O Marti. Science, **242**, 209 (1988).

[6] R J Colton, A Engel, J E Frommer, H E Gaub, A A Gewirth, R Guckenberger, J Rabe, W M Heckl & B Parkinson (Eds.). *Procedures in Scanning Probe Microscopies*. John Wiley and Sons (1998).

[7] C M Mate, R Erlandsson & G M McCelland. Journal of Vacuum Science & Technology A, **6**, 575 (1988).

[8] D F Evans, R Yang & G Lee. Journal of the Electrochemical Society, **135**, 394 (1998).

[9] *MultiMode™ SPM Instruction Manual*. Digital Instruments.

[10] O Marti, S Gould & P K Hansma. Review of Scientific Instruments, **59**, 836 (1988).

[11] C M Mate, R Erlandsson & G M McCelland. Journal of Vacuum Science & Technology A, **6**, 266 (1988).

[12] *A Practical Guide to Scanning Probe Microscopy*. Park Scientific Instruments.

[13] J Schneir, O Marti & G Remmers. Journal of Vacuum Science & Technology A, **6**, 283 (1988).

[14] Y N Moiseev, V M Mostepanenko & V I Panov. Physics Letters A, **132**, 354 (1988).

[15] U Landman. Abstract Papers of the American Chemical Society, **196**, 176 (1988).

[16] B Dwir, F Reinhardt, G Biasiol & E Kapon. Materials Science & Engineering B, **37**, 83 (1996).

[17] O Marti, S Gould & P K Hansma. Journal of Vacuum Science & Technology A, **6**, 287 (1988).

[18] C C Hsu, J B Xu & I H Wilson. Applied Physics Letters, **65**, 1394 (1994).

[19] T Fukui, J Ishizaki, S Hara, J Motohisa & H Hasegawa. Journal of Crystal Growth, **146**, 183 (1995).

[20] dopant reference!!

[21] B Dwir, F Reinhardt, G Biasiol & E Kapon. Applied Surface Science, **104**, 529 (1995).

[22] S Müller, J L Weyher, R Dian & W Jantz. Materials Science and Engineering B, **44**, 96 (1997).

Chapter 5

Experimental Considerations

5.1 Introduction

This chapter discusses factors that were considered during sample preparation and imaging. These involve the various aspects of sample preparation and sources of error.

Careful preparation of the sample wafers was essential to the successful metrology of device structures. After much 'trial and error' a cleaving technique was developed that produced consistently good quality cleaves that allowed the imaging of the oxide induced contrast on the device structures. A specially designed sample puck was also used in order to easily measure the cross-section of the samples.

Errors are unavoidable in any experimental work. It was therefore important to consider their source and determine what effect they had on any data acquired. Several sources of error were thought to exist during this research and are outlined in this chapter.

Section 5.2 discusses sample preparation, outlining various cleaving issues, cleaving technique and the mounting of the sample. Section 5.3 then discusses the various sources of error that may be encountered whilst carrying out AFM studies on the wafers.

5.2 Sample Preparation

5.2.1 Cleaving of the Wafer

In order for successful cross-sectional metrology to be performed, wafers must be cleaved to reveal a (110) surface that is as smooth as possible. Ideally, there

should be no steps (even at the atomic scale) over sufficiently large areas as to allow layer thicknesses and oxide step heights to be measured. Steps can be confused with and sometimes overwhelm those caused by oxide height differences. Although an extreme case, fig 5.1 illustrates the point. The steps here are, as is often observed, parallel to the surface and so also to the interface between epitaxial layers. Many of the steps are (3.4 ± 0.3) Å high, compatible with the lattice spacing for the (110) surface. This compares with differential oxide heights that could be < 1 Å.

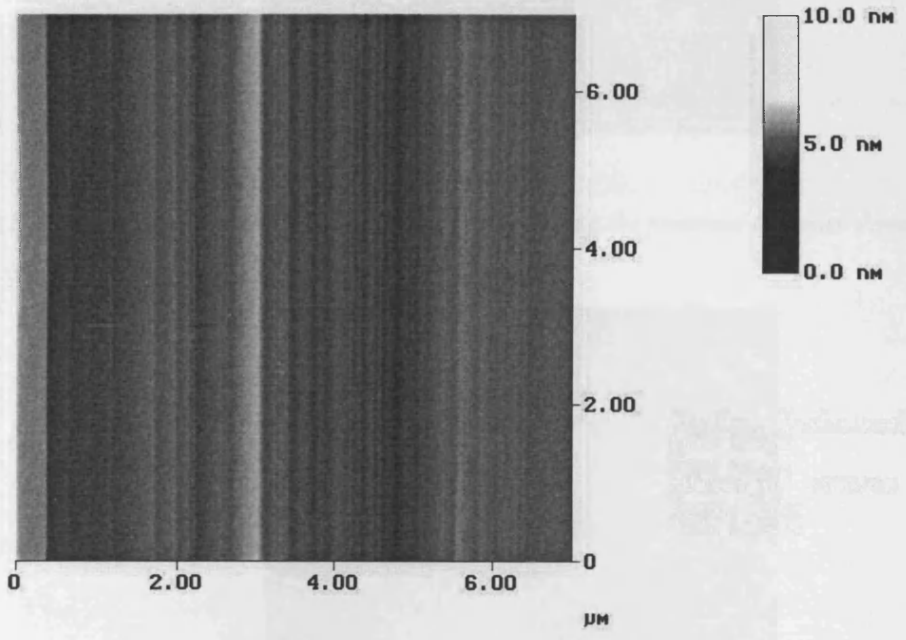

Fig 5.1 Cleave damage steps on the surface of a GaAs wafer.

Throughout this study, it was observed that the extent of cleave damage could vary considerably. Severe cleave damage is clearly illustrated in the optical photograph fig 5.2. Small scale cleave damage, such as that revealed in fig 5.1 could be present even on optically perfect surfaces such as that illustrated in fig 5.3.

Fig 5.2 Optical micrograph of a cleaved wafer illustrating the presence of cleave damage.

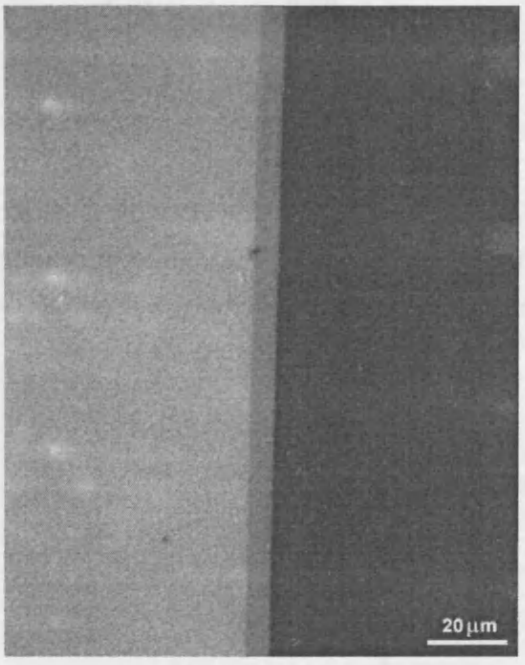

Fig 5.3 Optical micrograph of a cleaved wafer. The epitaxial layers can be seen as a dark band along the right hand edge of the wafer.

The nature of cleave damage could also depend on the sample structure itself. It was frequently observed that the direction of propagation of the steps could change

through different material layers, as illustrated in fig 5.4, or even stopped completely, as shown in fig 5.5.

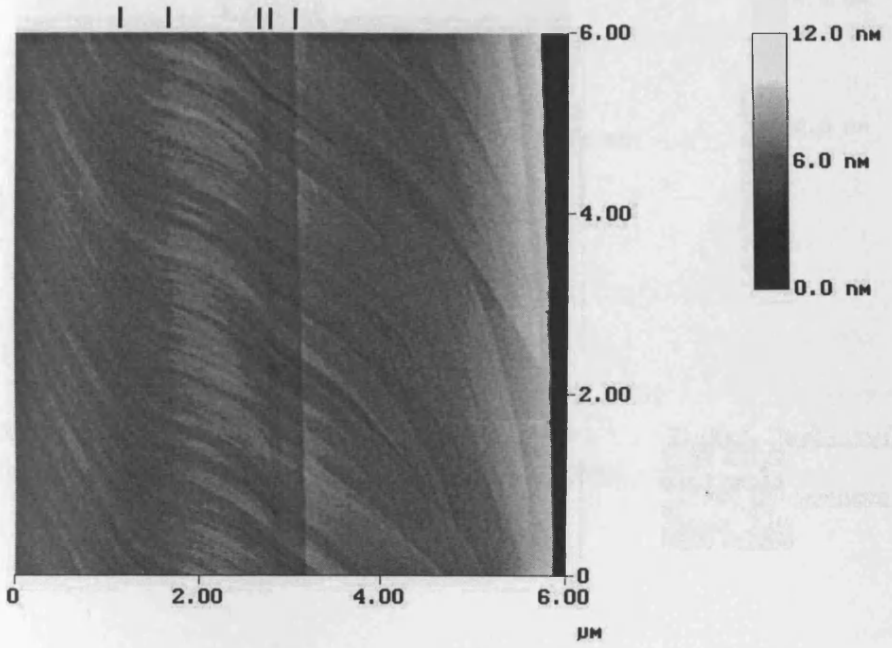

Fig 5.4 A topographic AFM image of a semiconductor heterostructure, with layer interfaces indicated by black lines. The direction of cleave damage propagation changes through the different layers.

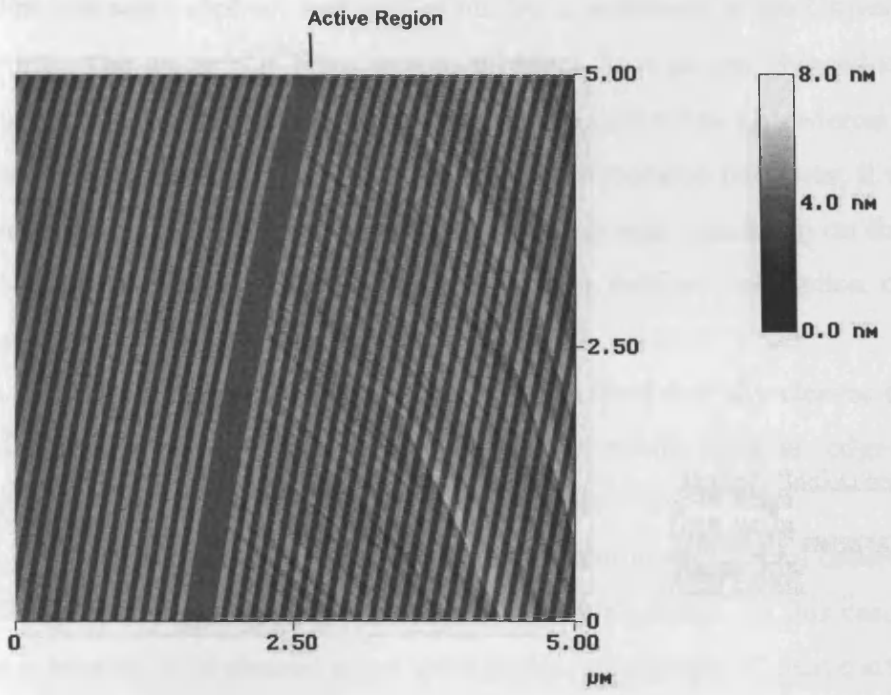

Fig 5.5 A topographic AFM image of a VCSEL structure. Cleave damage clearly propagates through the right hand mirror, but is stopped by the presence of different materials in the active region (dark band, centre).

5.2.2 Factors Affecting Cleave Quality

The quality of the cleave depended on the wafer thickness and the misorientation of the substrate. Thick wafers (~ 650 μm) required considerable force to cleave, and were particularly prone to suffering from large-scale cleave damage.

Any misorientation along the cleavage planes also had a detrimental effect on the cleave quality. A series of systematic cleaving experiments on GaAs wafers showed that the production of cleaved surfaces of a sufficiently high quality was 70 % for full thickness wafers with a 2° misorientation. However, this dropped significantly to only 10 % for full thickness wafers with a 10° misorientation.

Fortunately, the combined effects of wafer thickness and misorientation could be removed. Firstly, the wafers were lapped down to a thickness of ~ 150 μm, allowing the use of minimal force during the cleaving process. Standard lapping

techniques were applied, and carried out by a technician at the University of Wales Cardiff. The wafer was fixed to a metal block by wax, and thinned using a 22 μm abrasive pad. The thickness of the wafer was monitored using a Mercer 155 thickness gauge. Once the wafer had been thinned to the required thickness, it was recovered from the metal block by heating the wax. Any wax remaining on the sample was removed with warm trichloroethane. A more detailed description of the lapping procedure is included in appendix A.

Secondly, the geometry of the wafer dictated that any cleaves that revealed a (110) surface perpendicular to the major flat would have an edge that was not misorientated to the epitaxial surface of the wafer.

Thus the combination of thinning the sample and careful consideration of the wafer geometry significantly improved the cleave quality. In this case, it was found that at least 90 % of cleaved edges were completely devoid of cleave artefacts.

5.2.3 Wafer Cleaving

Wafers may be cleaved by simply applying light pressure to the edge with a scalpel blade. However, the production of a cleaved edge suitably devoid of damage as to allow successful metrology is not guaranteed. For this research, a simple method was developed that consistently provided excellent quality cleaves with as many as 90 % being suitable for metrology.

The first step was to crudely cleave the wafer down to an appropriate size, a rectangle roughly 10 × 5 mm. The wafer was then placed epitaxial layer side up on a piece of paper, on to which were drawn two parallel lines, 5 mm apart. Using an optical microscope as an aid, the short edge of the wafer was lined up with one of the parallel lines, so that the second was underneath and visible either side of the wafer.

With the aid of a small plastic rule and the line parallel to the short edge, two scribe marks were made on either side of the wafer, parallel to and 5 mm from the short edge. Care was taken so as to not make the scribe marks too long as these ultimately marked out the position of the cleaved edge and cut into the layers that were to be measured. It was also useful at this point to mark a small cross on the wafer in order to indicate where the position of the freshly cleaved edge would be.

The scribed wafer was then placed in a small clamp (cleave jig), as shown in fig 5.6. Again with the aid of an optical microscope, the scribe marks were lined up with the edges of the clamp, so that half of the wafer was overhanging the jaws. A light pressure was then applied in order to hold the wafer in place. The final step was to then tap the overhanging portion of the wafer lightly with the base of a pair of plastic tweezers. This force was sufficient enough to break the wafer, produce a freshly cleaved edge and a sample of dimensions ~ 5 mm × 5 mm.

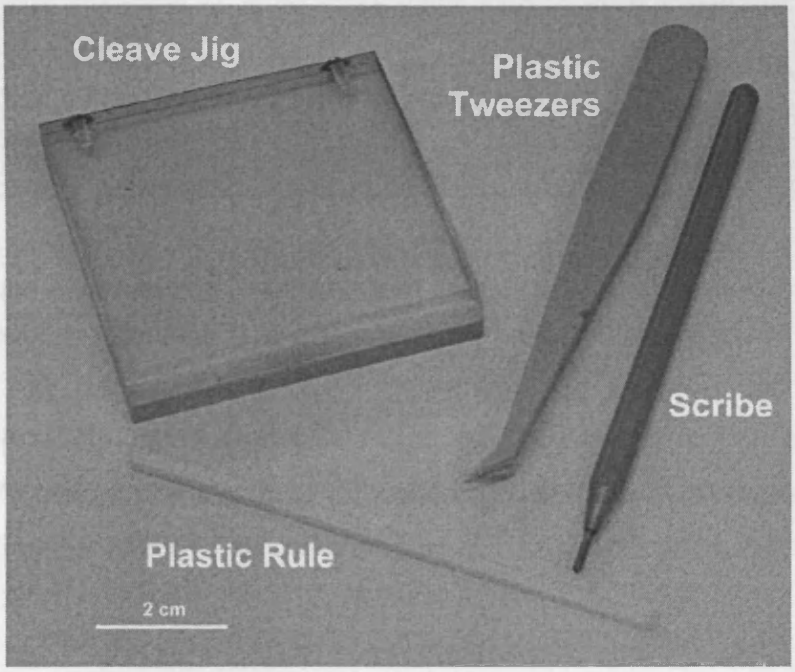

Fig 5.6 A digital camera image of the cleaving equipment. The implements used in the procedure were plastic as to minimise the damage to the wafer.

5.2.4 Mounting the Sample

Samples were mounted using a sample puck designed specifically for the cross-sectional examination of the of semiconductor wafers. The circular puck, 12 mm in diameter is shown in fig 5.7. It was fabricated from steel, allowing it to remain fixed to the magnetised region on top of the piezoelectric scanner. The wafer was placed resting against the central bar, with the surface to be examined facing upwards. The holding screw was then tightened, fixing the wafer into position. Care was taken as to not over-tighten the screw, particularly with brittle thinned wafers.

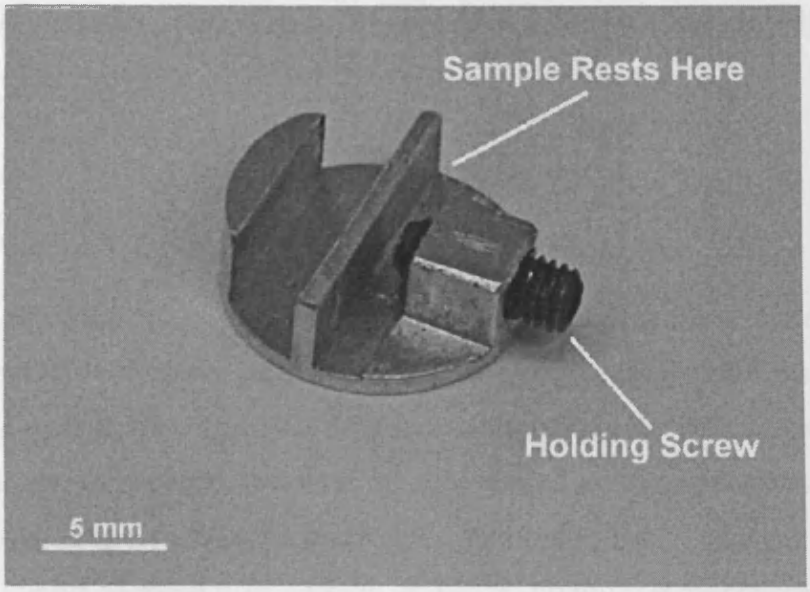

Fig 5.7 A digital camera image of the sample puck used for cross-sectional analysis. The wafer is propped up against the central bar and held in position by means of the holding screw.

In general, the time between the mounting of the sample and the acquisition of the first AFM image was less than 20 minutes. This is a relatively short sample preparation time compared to that for other techniques such as STM and TEM.

5.3 Sources of Error

Careful consideration of all the sources of error is essential in any research. This is particularly true in this case as metrology was carried out. If carefully calibrated, the AFM is capable of an accuracy of 2 - 3 % (as outlined in section 4.5.3). However, there were determined to be several other potential sources of error associated with this research. They are dealt with below.

5.3.1 Cleave Damage

A thorough treatment of the sources of cleave damage was presented above. Suffice to say that some cleave artefacts appeared along the interface between

heterolayers, and this compounded the difficulty in measuring accurate step heights. These cleave damage steps largely disappeared with the thinning of the sample wafers.

5.3.2 Scan Resolution

The selection of an appropriate scan size is an important factor in AFMs ability to resolve particularly narrow layers. For example, quantum wells will remain undetected in images where the scan step size is larger than the well thickness.

Similarly, an appropriate *z* limit must be chosen for the piezoelectric scanner if very small step heights are to be measured. Normally, the 'J' scanner (extensively used in this study) operates over a range of 440 V, giving it a maximum *z* deflection of ~ 6.25 μm [1]. The *z* direction control system uses a 16-bit digital to analogue converter (DAC). This means that the resolution of the scanner in the z direction is approximately 1 Å (6.25 μm divided by 2^{16} bits).

However, the voltage range of the scanner may be reduced to 55V, giving a maximum *z* deflection of 781 nm [1]. Since the resolution of the scanner remains unchanged at 2^{16} bits, the voltage reduction may achieve a spatial resolution of ~ 0.1 Å. This is comparable to the minimum cantilever deflection the optical lever method may detect [1].

5.3.3 The AFM Scanner

Calibration procedures for the piezoelectric scanner have already been described in section 4.5.3. However, the cleaved edges of the wafers examined were elevated to heights at least 7 mm above the scanner. During imaging, the AFM scanner traces an arc underneath the tip. The radius of this arc will be greater if samples are elevated significantly above the scanner, resulting in erroneous distance measurements. Investigation of a 2 μm pitch standard calibration sample revealed that a reduction of (10 ± 1) % is introduced at the heights involved in this research. Thus a 1 μm square image as measured by AFM actually imaged a 1.1 μm square region of the sample surface. This error was compensated for in all values quoted in

this work, by simply increasing them by 10 %. However, the error could not be corrected for in the images presented in this work, and are thus 10 % smaller than indicated.

Drift occurs in an image directly after a step change in applied voltage to the scanner. These changes are generally desired since they result in a change in scan size or a translation in a particular direction. However, for a short time after the voltage change, the image may suffer distortions as a result of one or more of the scanner nonlinearities outlined in section 4.5.2. Drift generally manifests itself in the slow-scan direction, since the scanner moves less quickly in this direction. Layer thicknesses were determined by taking the average heights along the slow-scan direction, where drift has an effect. Care was therefore taken to take the effects of drift into account before determining layer thicknesses.

5.3.4 Tip Convolution

The importance of careful consideration of the effects of tip convolution on the measured height profiles has been highlighted in section 4.4.2. Differences in oxidation rates for the different material layers of superlattice structures cause trench and ridge like features. These features are comparable in size to the AFM tip, a regime in which tip convolution effects become significant, resulting in erroneous measurements of oxide height and width.

A simple geometrical argument may be applied to model the tip interaction with oxide layers for a number of situations, as illustrated in fig. 5.8 [2]. In this case, the tip remains in constant contact with the surface, but the model is assumed to be valid for TappingMode™ AFM.

Three different situations are envisioned below for the interaction of the tip with the oxidising device structure. The first is the interaction with a trench feature, where the width of the trench is comparable to the radius of curvature of the tip. This situation is analogous to the one encountered when a quantum well is imaged. The second is the interaction with a ridge feature, namely a layer that is oxidising faster than the surrounding material. Finally, the interaction with a ridge feature of height greater than the radius of curvature of the tip is considered. This is relevant for layers

with relatively high oxidation rates, where the contact angle of the probe to the oxidising layer is of significance.

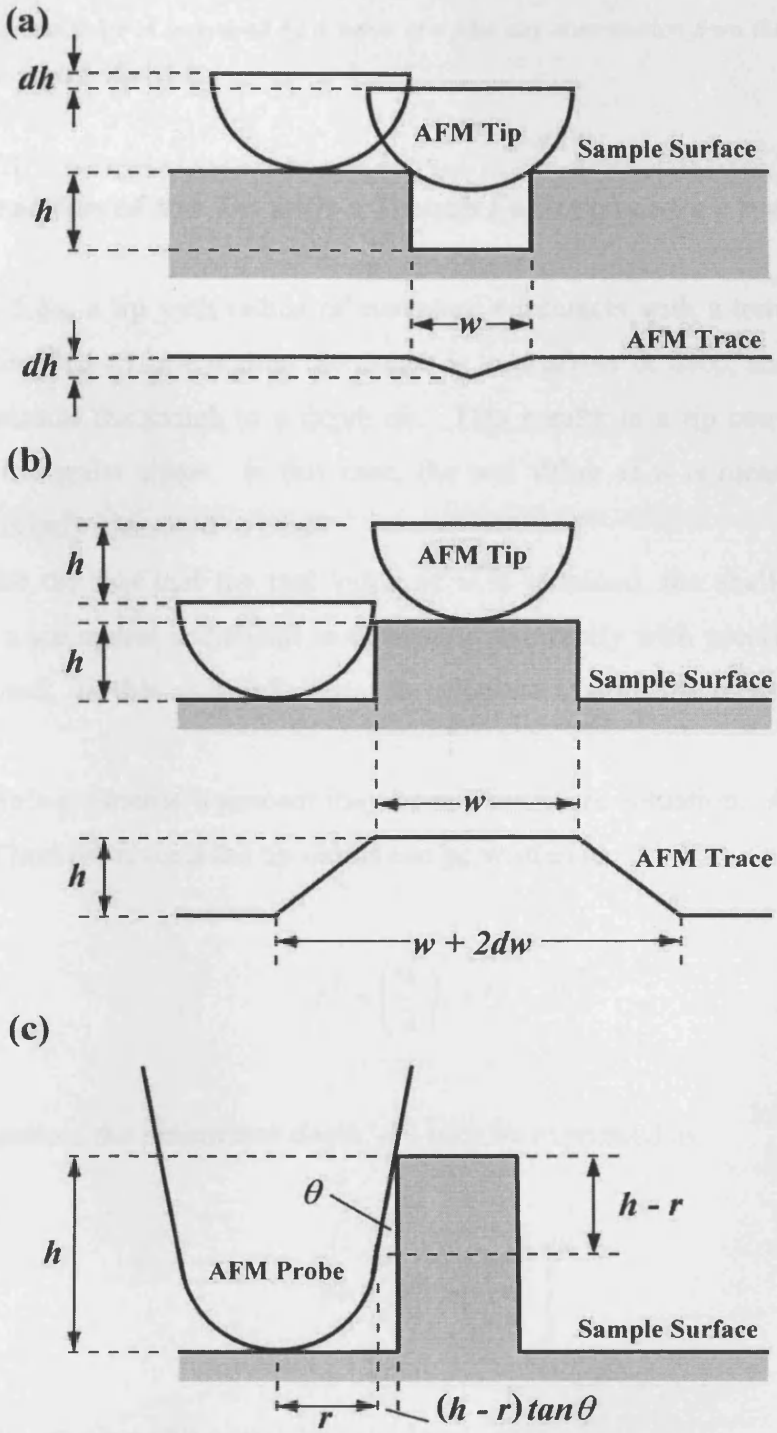

*Fig 5.8 A geometrical model of an AFM tip/probe with radius of curvature r scanning either a ridge or a trench of height/depth h and width w on an otherwise flat surface. **(a)** For a trench narrower than the radius of curvature of the tip, the apparent depth dh is independent of the real depth h; instead, it*

depends on w due to the partial penetration of the tip into the trench. The width however, is conserved. (b) For a ridge of any width, the apparent height h is the same as the real one, while the base width is increased by 2dw. (c) If the ridge is higher than the radius of curvature of the tip, then the contact angle θ between the probe and the side wall of the ridge contributes to the convolution. In this case, the base width of the ridge is increased by a value of r plus any contribution from the contact angle. The height however, is conserved.

5.3.4.1 Interaction of the Tip with a Trench Feature where r > w

In fig 5.8a, a tip with radius of curvature r interacts with a trench feature of width w and depth h. If $r > w$ then the trench is too narrow or deep, and the tip only partially penetrates the trench to a depth dh. This results in a tip convoluted AFM trace with a triangular shape. In this case, the real value of w is measured, but the trench depth is only measured to be dh.

Despite the fact that the real value of w is obtained, the shallow triangular shape of the trace makes it difficult to determine w directly with precision from any images obtained. In this case, it is better to calculate w from the measured value of dh.

A simple geometric argument may be applied to the situation. Application of Pythagoras' Theorem reveals the tip radius can be written as

$$r^2 = \left(\frac{w}{2}\right)^2 + (r - dh)^2 \qquad (5.1)$$

From this equation, the penetration depth, dh, may be expressed as

$$dh = \left[r^2 - \left(\frac{w}{2}\right)^2\right]^{1/2} \qquad (5.2)$$

from which the trench width w may be calculated.

5.3.4.2 Interaction of the Tip with a Ridge Feature where r < h

In fig 5.8b, the tip interacts with a ridge feature, of width w and height h. It is clear that in this case, the measured height will be the real one. However, the interaction of the probe with the side wall of the ridge results in a larger measurement of the ridge width. The resulting convoluted AFM trace is trapezoidal in shape.

Using a similar geometrical argument to the one applied above, the increase in width on one side of the ridge, dw, may be expressed as

$$\left[r^2 - (r-h)^2\right]^{1/2} \tag{5.3}$$

The total measured thickness of the ridge, $w + 2dw$ is therefore

$$w + 2\left[r^2 - (r-h)^2\right]^{1/2} \tag{5.4}$$

5.3.4.3 Interaction of the Tip with a Ridge Feature where r > h

A further complication to the above situation arises when the ridge feature is higher than the radius of curvature of the tip. It is necessary to consider the interaction of the ridge with the side of the probe, as illustrated in fig 5.8c. The convolution in this case results in a further increase of the measured width in addition to the increase outlined above.

If the probe makes an angle θ with the ridge wall, trigonometry reveals that the contribution to the measured ridge width is

$$2(h-r)\tan\theta \tag{5.5}$$

The contribution from the tip is its maximum, i.e. r. The total increase in measured ridge width is thus

$$2\left[r + (h-r)\tan\theta\right] \tag{5.6}$$

5.3.5 Growth Tolerances

Errors introduced in the composition and dimensions by the growth process are also of vital importance. An AlGaAs/GaAs test structure was manufactured by MOVPE at International Quantum Epitaxy Ltd. (IQE) to our design, that acted as a standard from which all other compositions were calibrated and oxide spreading thicknesses were determined (see chapter 6). The quoted tolerance for this sample was 2 - 3 % in both composition and layer thickness, together with an interfacial width between layers on the monolayer scale. While the determination of the periodicity of DBR layers in VCSEL devices may be achieved without reference to the test structure, it must be remembered that all other subsequent measurements on samples are subject to any errors inherent in the test structure.

5.4 References

[1] *MultiMode™ SPM Instruction Manual.* Digital Instruments.

[2] B Dwir, F Reinhardt, G Biasiol & E Kapon. Materials Science and Engineering B, **37**, 83 (1996).

Chapter 6

Studies of the $Al_xGa_{1-x}As$/GaAs Test Sample

6.1 Introduction

Before any successful metrology can be carried out on $Al_xGa_{1-x}As$/GaAs devices, it is necessary to determine how changes in oxide thickness and broadening can be related to layer compositions and thickness. To this end, a test sample was grown that allowed the determination of these quantities for a variety of $Al_xGa_{1-x}As$ compositions. This sample provided the data from which layer composition and thickness could be determined for real devices. In addition, detailed analysis was carried out on the results in order to determine the oxidation mechanisms for the various materials in the sample.

Section 6.2 outlines the design and motivation for the test sample. The results obtained from the sample are presented in section 6.3, while analysis of the data is conducted in section 6.4. Section 6.5 concludes the chapter by outlining how the results may be used as a tool for accurate layer composition and thickness determination.

6.2 Motivation and Design

The ultimate goal of this work is to produce data that will allow the determination of composition and layer thickness primarily for $Al_xGa_{1-x}As$/GaAs based optoelectronic devices. Preliminary experiments were carried out on real devices such as VCSELs that allowed the determination of typical oxidation rates and optimum scan conditions. These confirmed reports in the literature that the oxidation rate of layers was heavily dependent on composition, particularly for the higher aluminium concentrations in $Al_xGa_{1-x}As$ [1]. It was also noted here that the oxide

broadens, sometimes to thicknesses considerably wider than those of the original oxidising layer.

The problem with relying on measurements from real devices is that there is no control over thickness and compositions, so there is no consistency for comparison. Also, the surrounding materials tend to vary. The next step, therefore, was to design a sample that would enable the understanding of the oxidation and subsequently judge the potential of the technique for thickness and composition determination. The structure, specified in table 6.1, was designed to be fully representative of the range of $Al_xGa_{1-x}As$ materials, with more emphasis on higher values of x because of the oxidation behaviour.

The sample was fabricated at International Quantum Epitaxy Ltd. The epitaxial layers were grown on a (100) GaAs substrate misorientated 2° toward the (110) direction by MOVPE. The total thickness of the epitaxially grown layers was 7.05 μm. Each layer was *n* doped to 10^{18} cm^{-3}, typical of the doping levels found in III-V devices. The expected errors in layer thickness and composition are within ± 2 - 3 %. It should be noted that the structure is anisotropic, having the thicker, higher aluminium mole fraction layers near the surface. In addition, the amount of $Al_xGa_{1-x}As$ in the structure means that some degree of strain relaxation might be expected. However, the mismatch between $Al_xGa_{1-x}As$ and GaAs is so small (~ 0.16 %), that the effect on measured thicknesses should be insignificant.

For each composition, at least two different layer thicknesses were included, allowing the investigation of the dependence of oxidation rate on layer thickness. Larger GaAs spacers were placed between the layers containing most aluminium, since the effects of oxide broadening were expected to be significant for these layers. A larger range of thicknesses was also included for these layers, as the effect of layer thickness on oxidation rate was suspected to be more pronounced.

The main compromise during the design of this sample was the fact that the number of layers that could be included was limited. It would have been useful to have layers in the range x = 0.9 - 1.0, as it was known that oxidation rates increase significantly over this range. It should also be noted that oxide heights measured from this sample were with respect to GaAs. The effects of the oxidation of this material relative to the $Al_xGa_{1-x}As$ are expected to be insignificant for high but not for low aluminium concentrations.

Table 6.1 *Specification for the test sample. The GaAs substrate is at the bottom of the table and the layers are numbered in the order in which they were grown.*

Layer	Material	Composition (x)	Thickness (μm) ±10%
35	GaAs		0.50
34	$Al_xGa_{1-x}As$	1.00	0.40
33	GaAs		0.40
32	$Al_xGa_{1-x}As$	1.00	0.20
31	GaAs		0.40
30	$Al_xGa_{1-x}As$	1.00	0.10
29	GaAs		0.40
28	$Al_xGa_{1-x}As$	1.00	0.05
27	GaAs		0.40
26	$Al_xGa_{1-x}As$	0.90	0.20
25	GaAs		0.40
24	$Al_xGa_{1-x}As$	0.90	0.10
23	GaAs		0.20
22	$Al_xGa_{1-x}As$	0.90	0.05
21	GaAs		0.20
20	$Al_xGa_{1-x}As$	0.80	0.10
19	GaAs		0.20
18	$Al_xGa_{1-x}As$	0.80	0.05
17	GaAs		0.20
16	$Al_xGa_{1-x}As$	0.70	0.10
15	GaAs		0.20
14	$Al_xGa_{1-x}As$	0.70	0.05
13	GaAs		0.20
12	$Al_xGa_{1-x}As$	0.50	0.10
11	GaAs		0.20
10	$Al_xGa_{1-x}As$	0.50	0.05
09	GaAs		0.20
08	$Al_xGa_{1-x}As$	0.30	0.10
07	GaAs		0.20
06	$Al_xGa_{1-x}As$	0.30	0.05
05	GaAs		0.20
04	$Al_xGa_{1-x}As$	0.10	0.10
03	GaAs		0.20
02	$Al_xGa_{1-x}As$	0.10	0.05
01	GaAs		0.50
S/S	GaAs		650.00

6.3 Measurements of the $Al_xGa_{1-x}As/GaAs$ Test Structure

6.3.1 Nature of Measurements

The test structure, and subsequent device structures were probed using TappingMode™ AFM. Due to the reported porous nature of $Al_xGa_{1-x}As$ oxides, it was determined that this method was less likely to damage the oxidising surface than contact mode. In addition, it has been shown that while contact and TappingMode™ AFM lead qualitatively to the same result when studying these samples [1], TappingMode™ has a lower noise level. This reduction in nose level is an obvious advantage, particularly when trying to resolve particularly small oxide step heights.

In total, 10 individual cleaves were measured, allowing a detailed statistical analysis of the results to be carried out in order to determine the errors associated with the measurements. This analysis is presented in appendix B. Measurements of the oxide height h and width w, were obtained for each $Al_xGa_{1-x}As$ layer at various times after cleaving by using the profiles extracted from the topographic measurements of the sample. These profiles were obtained from the data by taking line averages along directions parallel to the layer interfaces using the AFM software. This was used to eliminate as much as possible any spurious errors and variations that may have occurred during the acquisition of a single scan line. Measurable changes in oxide height were observed over the period of one scan for the fast oxidising AlAs layers. Line averages were therefore conducted over half a scan as a compromise between scan line errors and oxide growth. The line averages were thus more representative of the profile. Care was taken during this procedure, as failure to take the line average perfectly parallel to the layer interfaces would result in artificial broadening of the oxide layers.

The oxide height and width were not intrinsically difficult to measure from the profiles obtained. However, the nature of the oxide growth meant obtaining accurate and representative values was not straightforward. The oxide profiles could take a number of different shapes as discussed later in section 6.4.1. For consistency, it was decided that in general, oxide heights should be taken from the level of the surrounding GaAs to the highest line averaged point of the oxide. Errors in measuring oxide heights were governed by the height fluctuations in the surrounding

oxidising GaAs, and were determined to be (± 0.5 Å). In the case of the $Al_{0.3}Ga_{0.7}As$ layers, where oxide heights of the order of 0.2 nm were observed, the error was therefore as much as 25 %. This is particularly significant, as the oxide heights were ultimately used as the basis from which layer compositions of unknown layers could be determined.

Oxide width measurements proved to be more problematical. Lateral growth of the oxide meant that the width of the oxidising layer could not be determined by direct measurement of the width of the oxide. Several different methods of width measurement were therefore considered. Each of these methods is illustrated below, and the pros and cons of each outlined.

6.3.1.1 Oxide Base Width Measurements

Of the possible ways to measure the layer thickness, direct measurement of the base width of the oxide growth is (see fig 6.1) the easiest to conduct. However, oxide layers grow laterally, as well as vertically, resulting in a 'broadening' of the original layer thickness, and an apparent change in position of the interface between adjacent layers. Therefore, in order to calculate the actual layer width from the width of the broadened oxide, the broadening of the oxide must be quantitatively understood. Given the complex oxidation behaviour of the materials, this is not a trivial matter.

It is also important to take the effect of tip convolution into account, as this too contributes to the broadening effect, albeit artificially. The extent of this artificial broadening is related to the tip radius, which can vary as the tip wears. It is therefore impossible to predict with any degree of certainty the effect of tip convolution. However, assuming that the tip radius will be 20 nm for the worst case, a maximum contribution to the broadening of the oxide layer by tip convolution may thus be calculated from equation 5.4. This gives a broadening value of 12.5 nm for an oxide layer 1 nm thick.

Finally, the profile of the oxide layer must be considered. For the profile shown in fig 6.1, the height change between the oxide layer and GaAs is sufficiently abrupt to allow the base width measurement to be carried out with confidence. However, errors caused by fluctuations in the surrounding GaAs oxide become significant as the aluminium concentrations and hence oxidation heights decrease.

This leads to a decrease in confidence of the layer width measurement. It was estimated that while the width of the oxide layers of AlAs could be determined to within 4 nm, that of $Al_{0.3}Ga_{0.7}As$ could at best be determined within 8 nm.

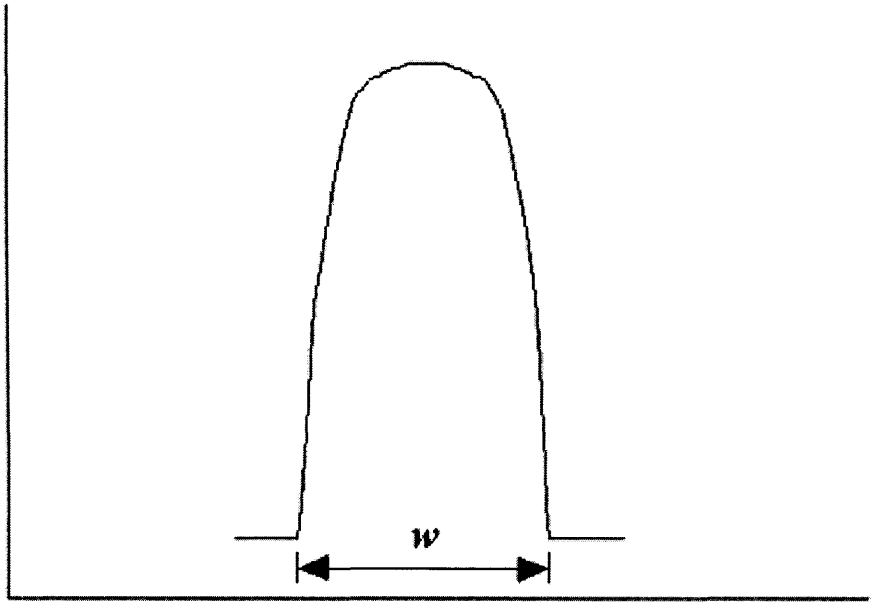

Fig 6.1 Typical shape for an oxide profile growing on an $Al_xGa_{1-x}As$ layer. Oxide base width values were measured as indicated.

6.3.1.2 Full Width Half Maximum Measurements

Although slightly more complicated and time consuming to measure, the full width half maximum (FWHM) was considered as an initial alternative to base width measurements. These measurements were carried out as indicated in fig 6.2. However, as discussed later, a number of different oxide profiles were observed each relating to a different stage of oxidation. It was difficult, therefore, to be consistent with the definition of this quantity throughout the sample. In particular, the FWHM measurement became largely useless when attempting to draw comparisons between layers.

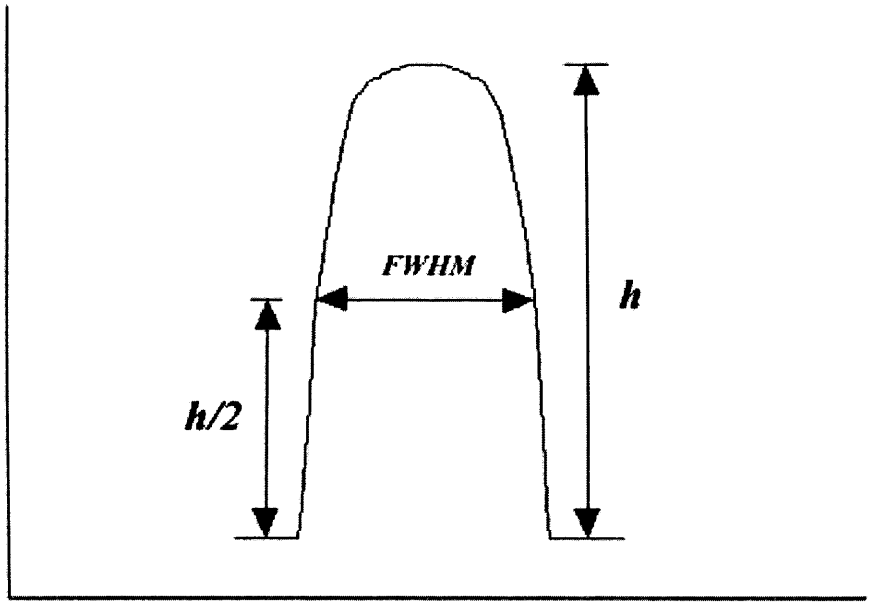

Fig 6.2 Typical shape for an oxide profile growing on an $Al_xGa_{1-x}As$ layer. Full width half maximum values were measured as indicated.

6.3.1.3 Profile Differential Measurements

The final type of measurement to be considered was that of measuring the differential of the oxide profile, as shown in fig 6.3. The process of obtaining the differential is easily performed by the AFM data analysis software. The measured distance is that between the two differential peaks as indicated in fig 6.3. These peaks need to be clearly defined in order to ensure confidence in these measurements, however.

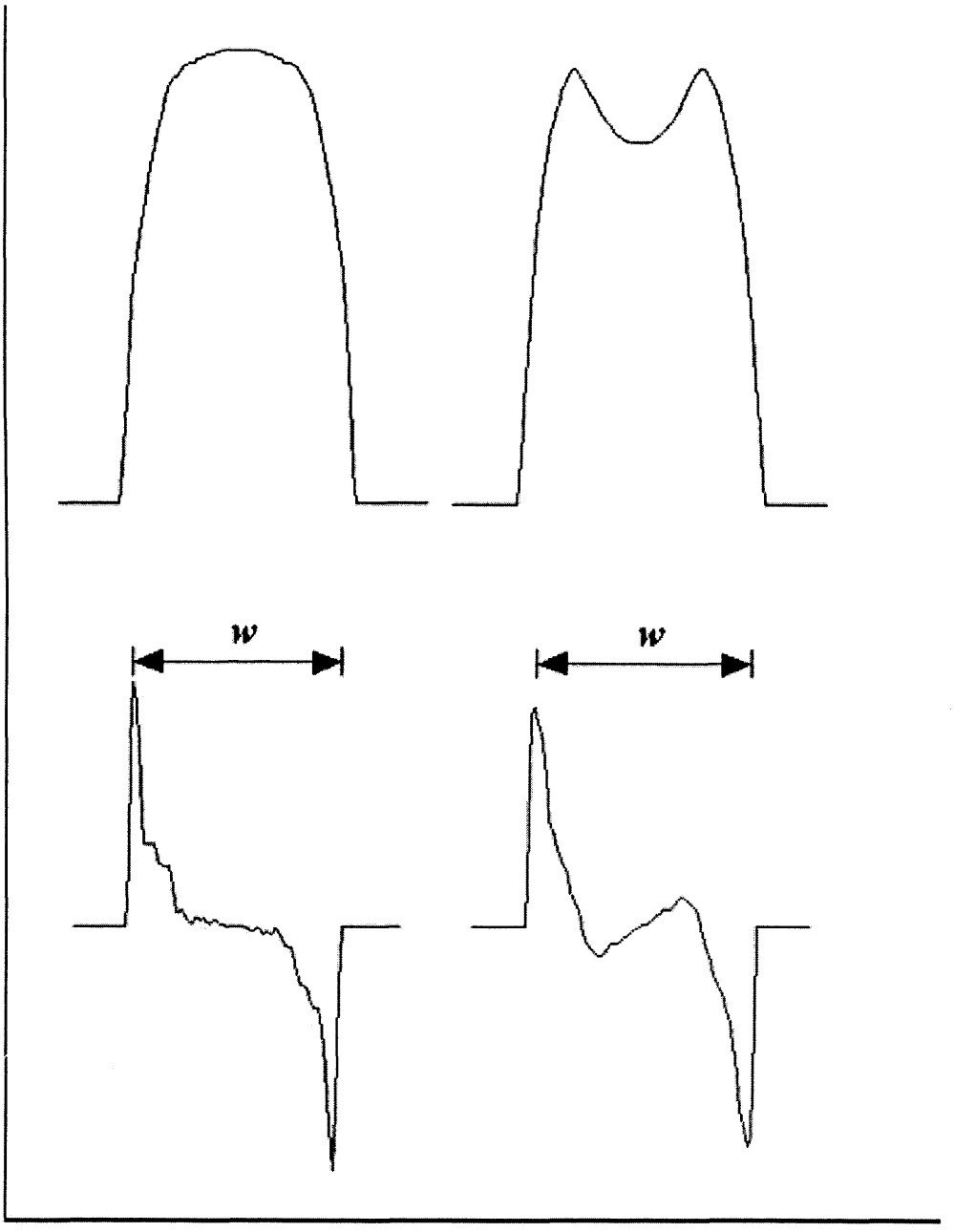

Fig 6.3 Two typical profile shapes for oxide growing on an $Al_xGa_{1-x}As$ layer. The differential of each layer is plotted below. The layer width is measured as indicated.

6.3.1.4 A Comparison of Measurement Techniques

In order to compare the three different approaches to width measurements, the time evolution of each of the quantities is plotted for the 400 nm AlAs layer in fig 6.4 below. The oxide heights varied from 80 nm to 170 nm. It is clear which is the best choice. The base width measurement is affected by oxide broadening and tip convolution effects, which result in larger than expected values. FWHM values are always less than the real value, although approaches it as the oxide height increases. This method is likely to be difficult, however, as complications arise due to the evolution of different oxide profiles. Calculation of the differential of the profile appears to overcome these limitations, however. This method allows the direct measurement of layer widths, removing any ambiguities that may be present with the first two methods.

Fig 6.5 compares oxide profile and associated derivative at selected times for the 400 nm AlAs layer. The evolution of the oxide profile and its associated derivative over time is clearly illustrated. It should be noted that the change in profile shape and height do not effect the value of the derivative peak-peak width. For each layer of the test structure, the distance between the peaks was found within errors, to equal layer width. Intuitively, this makes sense, as the largest change in oxidation rate occurs over the interface between the two layers. The error in this case was assumed to be the standard deviation of the spread in the peak to peak distance. This had a value of 3 nm at most.

In addition, for each layer, the height of these peaks was large enough to ensure confidence in the measurement. Furthermore, the measured value does not change significantly for extended periods of oxidation and appears unaffected by tip convolution. In summary, this method of layer width measurement is generally applicable and is thus a very powerful technique.

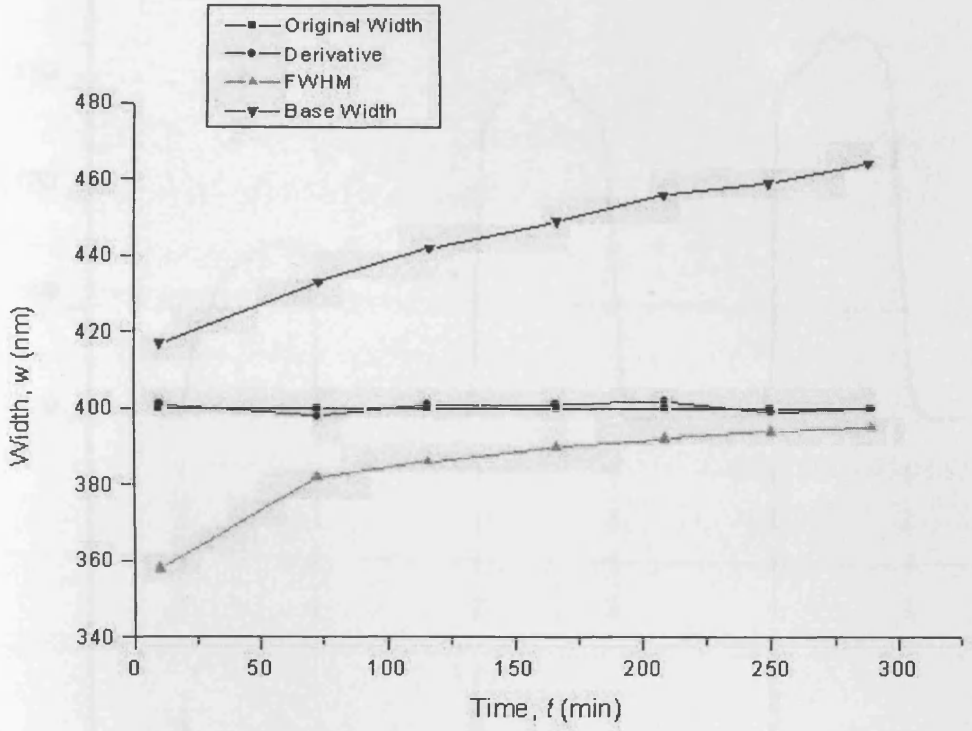

Fig 6.4 The time evolution of base width, FWHM and differential measurements for the 400 nm AlAs layer.

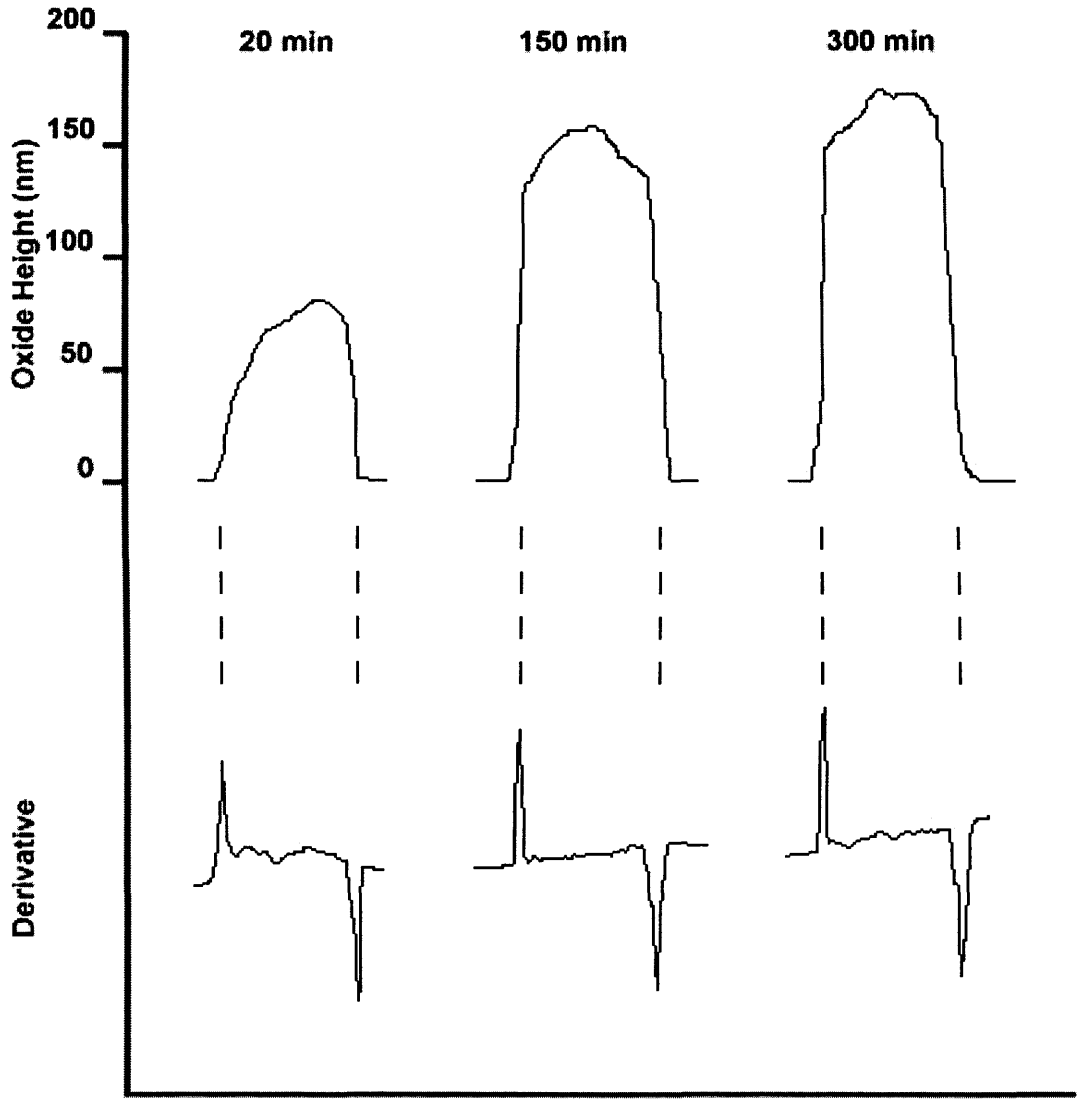

Fig 6.5 A comparison of oxide profile and associated derivative at selected times for the 400 nm AlAs layer.

6.3.2 General Observations

High quality AFM images could be acquired from the sample for up to 4 days after cleaving. After this time, the oxide growth on the AlAs layers (particularly the fast oxidising 400 nm thick layer) began to disintegrate, producing debris within the immediate vicinity of the epitaxial layers. For extended oxidation times of more than a week, delamination of the AlAs layers caused severe damage to the sample, rendering it unsuitable for further study.

Cleaving the full-thickness wafer (650 μm) produced the usual multiple small height cleave damage lines. However in addition, very high steps (0.5 - 1 μm) also occasionally formed at the interfaces between GaAs and the thicker, higher mole fraction $Al_xGa_{1-x}As$ layers. Such large steps are noted here because they were not observed on any other device investigated during this study. Their origin is uncertain but may lie in either the strain between the layers (~ 0.16 % between GaAs and AlAs), strain relaxation producing dislocations or to chemical differences. Although these artefacts largely disappeared by thinning the wafer down to 150 μm, some large steps were still occasionally observed.

During data acquisition, it was found that the continuous scanning of the same area of a sample is not desirable. A comparison of plots of oxide heights versus time acquired by continuously scanning the same region, and those obtained by scanning different areas of the sample were made. This revealed that the oxidation of the sample was inhibited by as much as 5 % for AlAs layers over the period of 18 hours. However, this retardation becomes less evident for layers with lower aluminium concentrations, dropping to 2 % for the $Al_{0.9}Ga_{0.1}As$ layers, and becoming unmeasurable for layers with smaller aluminium concentrations. No difference was observed when a doped tip was substituted for an undoped one.

A number of factors could have contributed to this retardation effect. Firstly, the higher mole fraction layers could be more porous and therefore more susceptible to being worn down by continuous tip action. This is evidenced by the disintegration of the AlAs oxide layers after 4 days. Finally, higher lateral forces will be experienced as the tip scans over higher oxide layers, causing more wearing to occur. Unfortunately, it was not practical to continuously change the scan region, so the effect was reduced by prolonging the time between successive scans.

6.3.3 Measurements of $Al_{0.1}Ga_{0.9}As$ Layers

The only layers not apparent on the topographical images after 20 minutes of oxidation were the ones composed of $Al_{0.1}Ga_{0.9}As$. This is perhaps not surprising, since the $Al_{0.3}Ga_{0.7}As$ oxide layers only achieved a height of 2.4 Å by this time. The oxidation rate of $Al_{0.1}Ga_{0.9}As$ is clearly very similar to that of the slowly oxidising

GaAs, although the presence of these layers was evident after a period of 3 days, with a step height of 0.5 Å.

Despite the lack of any initial topographic contrast for the $Al_{0.1}Ga_{0.9}As$ layers, their presence was confirmed by the use of phase imaging (outlined in section 4.3.4). Fig 6.6 shows the height and corresponding phase image of the $Al_{0.1}Ga_{0.9}As$ layer region. The presence of two additional layers in the phase image is immediately obvious. The origin of the phase contrast observed is thought to lie in differences in the 'stickiness' of oxidised $Al_{0.1}Ga_{0.9}As$ and GaAs.

The right hand layer shows the effect of relatively large height variations on the phase data. The dark and light edges are caused by the bigger error signal generated by the tip surmounting the step. However, the phase difference recorded for the centre of the layer is due to the properties of the oxide only.

It should be noted that the measured phase differences are not reproducible. This is due to the nature of the phase measurement. The probe is forced to tap the surface relatively hard, and the measured phase differences are due to a combination of the selected scanning conditions and the resonant frequency of the cantilever. Since the resonant frequency of each cantilever is different, then phase differences vary each time a different tip is used.

Phase difference values are thus largely meaningless, but they do reveal a trend for increasing phase difference with increasing x. For example, the measured phase difference between both $Al_{0.1}Ga_{0.9}As$ layers and GaAs for the layers in fig 6.6 is 0.448°. For $Al_{0.3}Ga_{0.7}As$ it is 0.761° and for $Al_{0.5}Ga_{0.5}As$ it is 1.129°. The phase differences are thus clearly dependent on the layer composition, but not the layer thickness.

Phase differences between oxide layers have not been reported before, and may be useful for the detection of materials with particularly low (or very similar) oxidation rates, including phosphide based materials. In addition, it should be noted that the absence of significant oxidation (and hence topographical features) aids the determination of interface position (and hence layer width determination), since the surface oxide will not have broadened. The layer widths are measured to be (51 ± 4) nm and (98 ± 4) nm. The error associated with these measurements derives from the spread over data collected from ten separate samples.

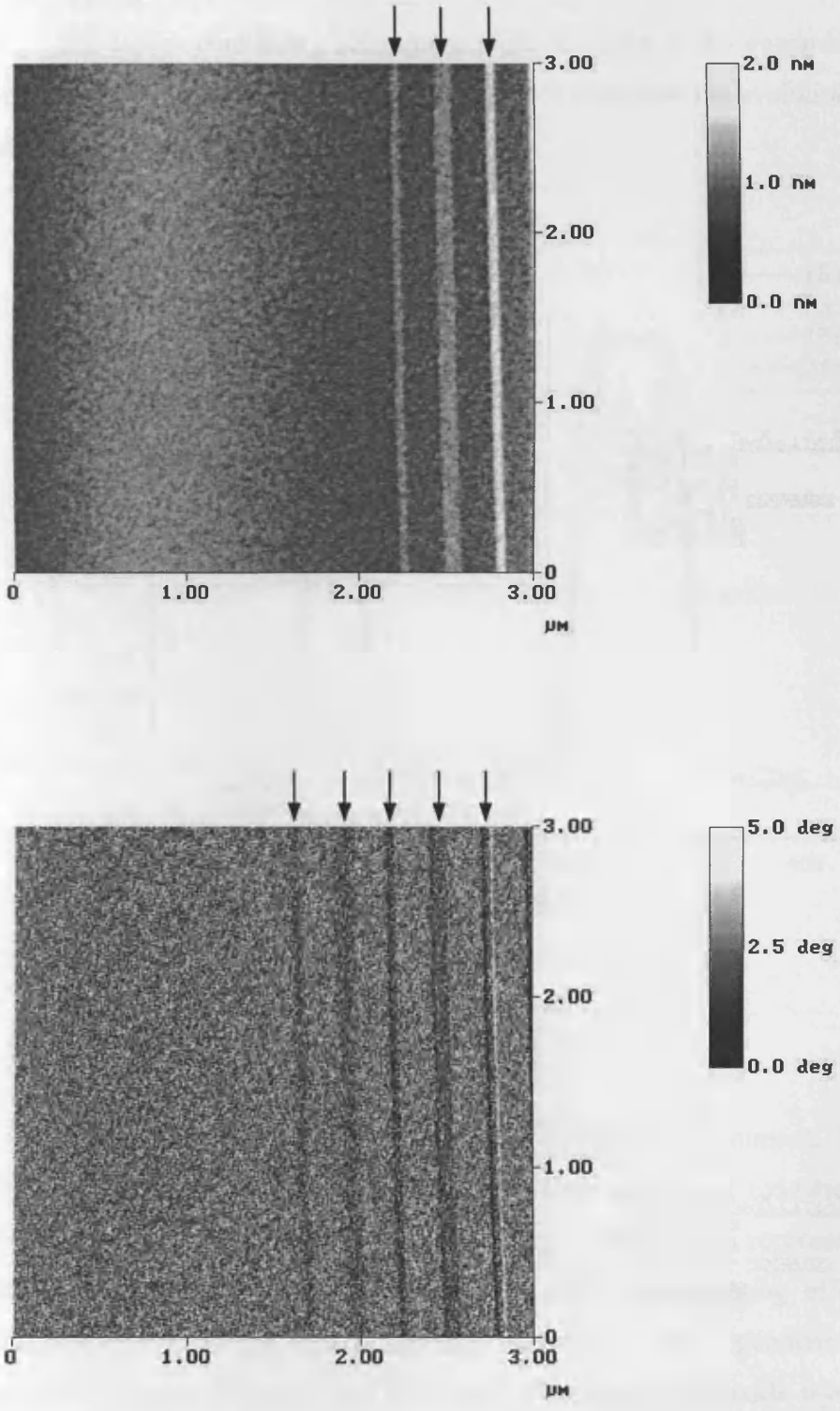

Fig 6.6 AFM topographic (above) and phase (below) images demonstrating the ability of phase imaging to locate the positions of two $Al_{0.1}Ga_{0.9}As$ layers that cannot be detected topographically. Arrows indicate the positions of the layers. Arrows indicate the position of layers.

6.3.4 Measurements of $Al_{0.3}Ga_{0.7}As$ - $Al_{0.8}Ga_{0.2}As$ Layers

The layers containing aluminium mole fractions of 0.3 upwards could all be readily detected by topographic contrast. Fig 6.7 illustrates the evolution of the oxide profile of the $Al_{0.7}Ga_{0.3}As$ layers with time.

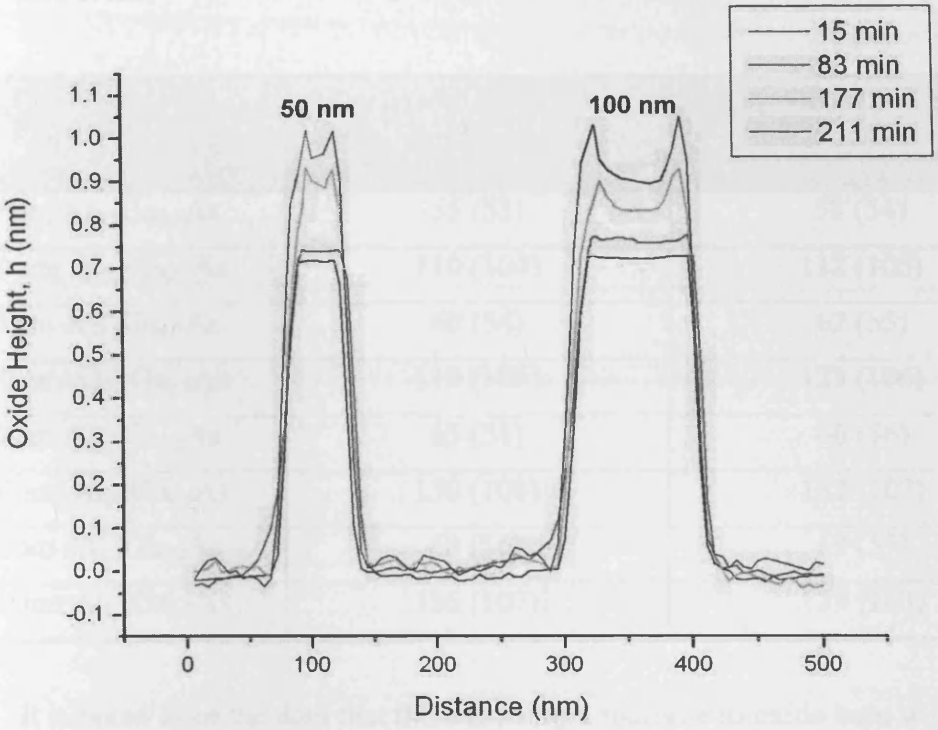

Fig 6.7 The evolution of the oxide profiles of the $Al_{0.7}Ga_{0.3}As$ layers with time.

The evolution of the oxide layer profile is worthy of comment. The top of the 100 nm layer in particular became crowned after prolonged oxidation. With the exception of the $Al_{0.3}Ga_{0.7}As$ layers, this change in shape was representative of that observed for the other layers in this range, with the crowning effect becoming apparent earlier for increased aluminium concentration. No dependence of the shape transition on oxide thickness was observed. The tops of the oxide layers associated with $Al_{0.3}Ga_{0.7}As$ remained flat throughout the lifetime of the sample.

As expected, the base widths of the oxide growth were larger than the layer width in each case. In addition, the widths are larger than would be expected purely by tip convolution (as calculated by equation 5.4), even with a tip radius of 20 nm.

This indicates that lateral growth of the oxide has taken place. These widths are summarised in table 6.2 below for times of 15 and 1650 mins after cleaving. Figures in brackets indicate the expected oxide base width with a tip radius of 20 nm.

Table 6.2 *Oxide base width measurements for the layers in the range x = 0.3 - 0.8 after 15 and 1650 mins of oxidation. The number in brackets indicates the expected width due to tip convolution for a tip of radius 20 nm.*

Layer	Base Width (nm) ± 4 nm 15 mins after cleave	Base Width (nm) ± 4 nm 1650 mins after cleave
50 nm $Al_{0.3}Ga_{0.7}As$	55 (53)	58 (54)
100 nm $Al_{0.3}Ga_{0.7}As$	110 (104)	112 (105)
50 nm $Al_{0.5}Ga_{0.5}As$	60 (54)	62 (55)
100 nm $Al_{0.5}Ga_{0.5}As$	119 (105)	123 (106)
50 nm $Al_{0.7}Ga_{0.3}As$	65 (55)	66 (56)
100 nm $Al_{0.7}Ga_{0.3}As$	130 (106)	132 (107)
50 nm $Al_{0.8}Ga_{0.2}As$	68 (56)	69 (57)
100 nm $Al_{0.8}Ga_{0.2}As$	136 (107)	139 (108)

It is noted from the data that there is a slight increase in oxide base width with time. It is possible that these increases are due to an increase in oxide base width due to the oxidation process. However, the increases are within the regime where they could be due to tip convolution. These effects should therefore not be ruled out, and it is possible that the increases are entirely due to the interaction of the tip with the layer.

Fig 6.8 shows plots of the mean oxide height difference between oxidised $Al_xGa_{1-x}As$ and GaAs as a function of oxidation time for these layers.

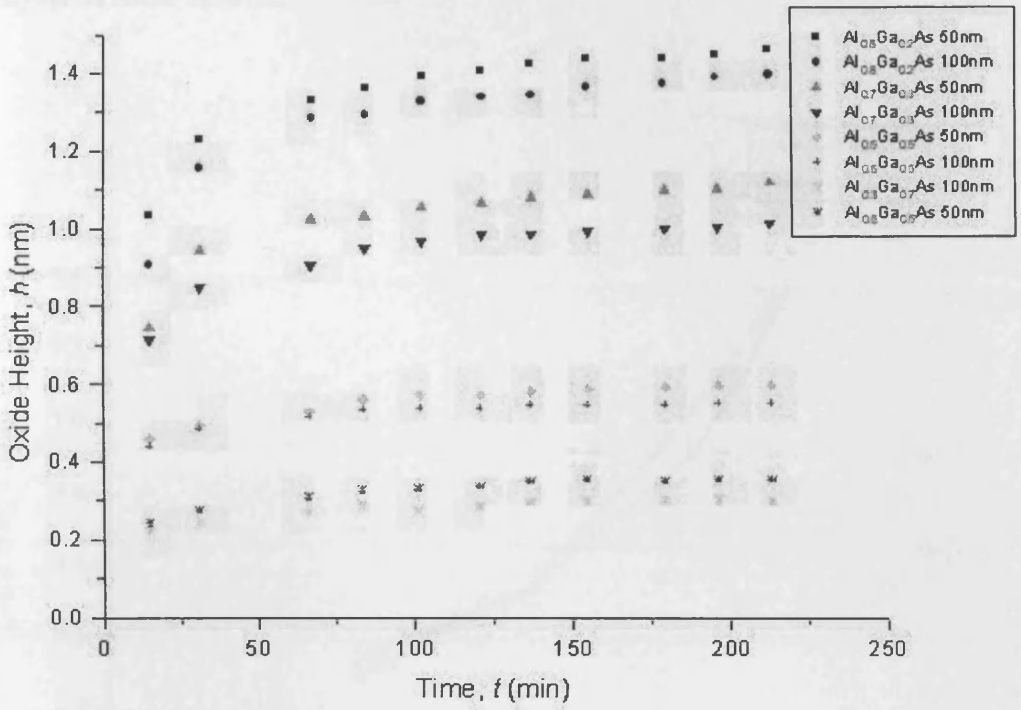

***Fig 6.8** A plot of oxide height difference versus time for the layers within the range x = 0.3 - 0.8.*

This plot clearly illustrates the effects of aluminium concentration and layer thickness on the oxidation rate. For all layers, the oxidation rate increases with aluminium concentration and layer thickness, the former being the most important parameter. After rapid initial oxidation, the oxidation slows down considerably, as suggested in the literature [2]. When the effects of GaAs oxidation relative to $Al_xGa_{1-x}As$ are taken into account, these results compare well to those obtained by other groups [1], presented in section 3.2.3.2. Although there were differences in composition and layer thickness between the data here and in section 3.2.3.2, reasonable comparisons could be made. An oxide height of roughly 0.6 nm for a 150 nm wide $Al_{0.5}Ga_{0.5}As$ layer after 100 min compared very favourably with the result presented here of (0.58 ± 0.08) nm for the 100 nm thick layer.

Fig 6.9 plots oxide heights against aluminium mole fraction 2 hours after cleaving for layers within this range. Here, the two layer thicknesses are plotted separately, showing the relationship between oxide height and aluminium mole

fraction for fixed layer width. This relationship shall be investigated further during the analysis of these results.

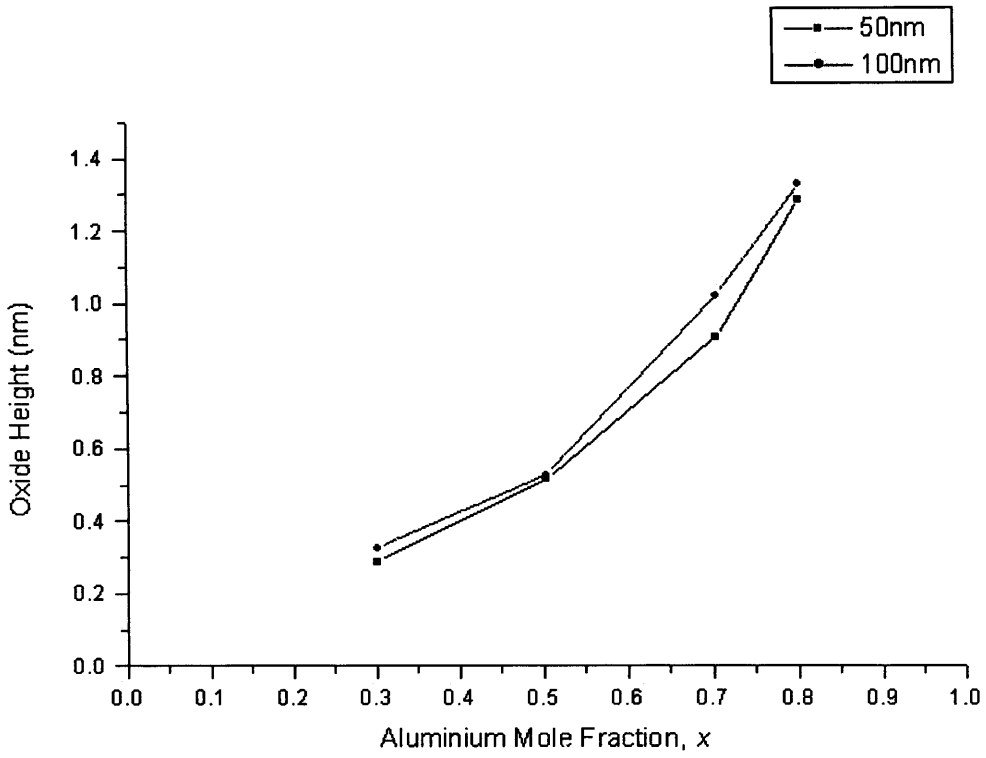

Fig 6.9 A plot of oxide height against aluminium mole fraction 2 hours after cleaving for layers within the range x = 0.3 – 0.8.

6.3.5 Measurements of $Al_{0.9}Ga_{0.1}As$ Layers

Three layers of this material were grown, of thickness 50 nm, 100 nm and 200 nm. The evolution of the oxide profiles of the two narrower layers with time is illustrated in fig 6.10. The profile of the 200 nm wide layer has been omitted to preserve clarity.

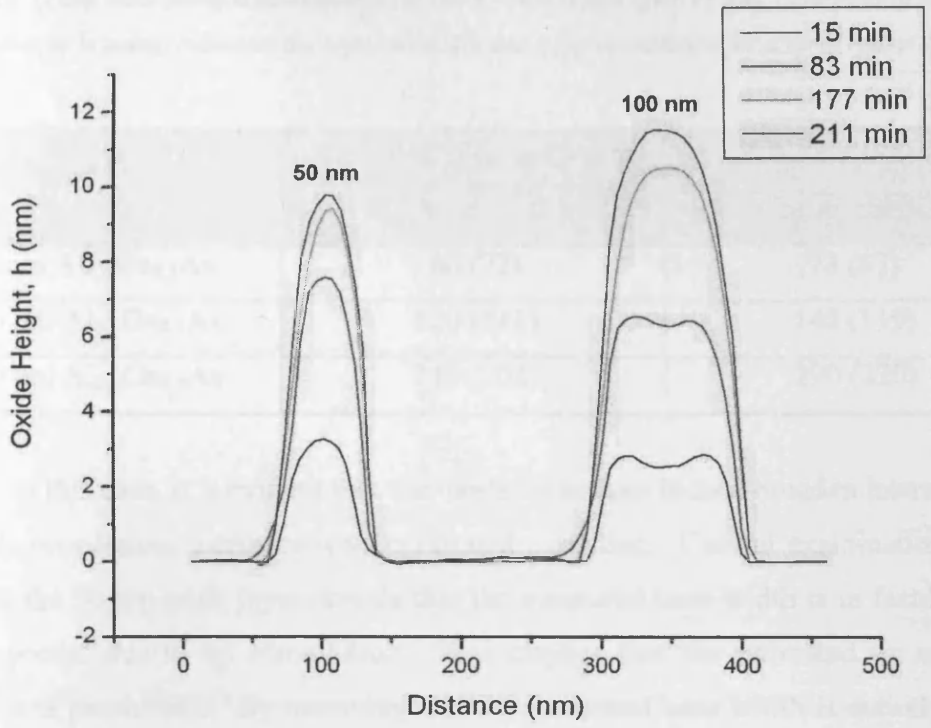

Fig 6.10 *The evolution of the oxide profiles of the 50 and 100 nm $Al_{0.9}Ga_{0.1}As$ layers with time.*

The oxide profile of the 50 nm layer appeared to stay the same with increasing oxidation, having a rounded top. The 100 nm thick layer, however, displayed the same crowning effect observed previously. This was also true of the 200 nm layer. After an oxidation time of 211 min, the crowned region of both the 100 nm and 200 nm wide layers, it was estimated to be (30 ± 6) nm for each interface. As oxidation continued, the width of the crowned regions increased until they eventually merged together producing a round-topped profile, similar to that of the 50 nm layer. This width increase occurred more quickly for the 100 nm wide layers. Combined with the fact that the layer was narrower, this resulted in the oxide achieving a round-topped profile more quickly than that of the 200 nm thick layer.

Lateral broadening of the oxide was also evident for these layers. These widths are summarised in table 6.3 below for times of 15 and 1650 min after cleaving. Figures in brackets indicate the expected oxide base width with a tip radius of 20 nm.

Table 6.3 Oxide base width measurements for the x = 0.9 layers after 15 and 1650 mins of oxidation. The number in brackets indicates the expected width due to tip convolution for a tip of radius 20 nm.

Layer	Base Width (nm) ± 4 nm 15 mins after cleave	Base Width (nm) ± 4 nm 1650 mins after cleave
50 nm $Al_{0.9}Ga_{0.1}As$	60 (72)	73 (67)
100 nm $Al_{0.9}Ga_{0.1}As$	120 (111)	145 (119)
200 nm $Al_{0.9}Ga_{0.1}As$	240 (208)	290 (220)

In this case, it is evident that the oxide layer does indeed broaden laterally, and that this broadening increases with continued oxidation. Careful examination of the data for the 50 nm wide layer reveals that the measured base width is in fact less that that expected due to tip convolution. This implies that the estimated tip radius of 20 nm was pessimistic. By assuming that the measured base width is entirely due to tip convolution, a tip radius of 15 nm is obtained. The tip radius can thus be expected to be less than this value.

Fig 6.11 shows plots of the mean oxide height difference between oxidised $Al_{0.9}Ga_{0.1}As$ and GaAs as a function of time for these layers. It shows clearly that for this material, the oxidation rate of the narrower layers is initially faster than that of the wider ones. However, after prolonged oxidation, the oxide on the wider layers became thicker, as expected, indicating a faster reduction in oxidation rate for the thin layers. On closer inspection, this can be observed in the evolution of the oxide profiles as illustrated in fig 6.10.

This behaviour differs from that of the lower mole fraction layers and the subsequent data obtained for the AlAs layers. It is clearly related to the crowing effect and will be examined in more detail during the analysis presented later in this chapter.

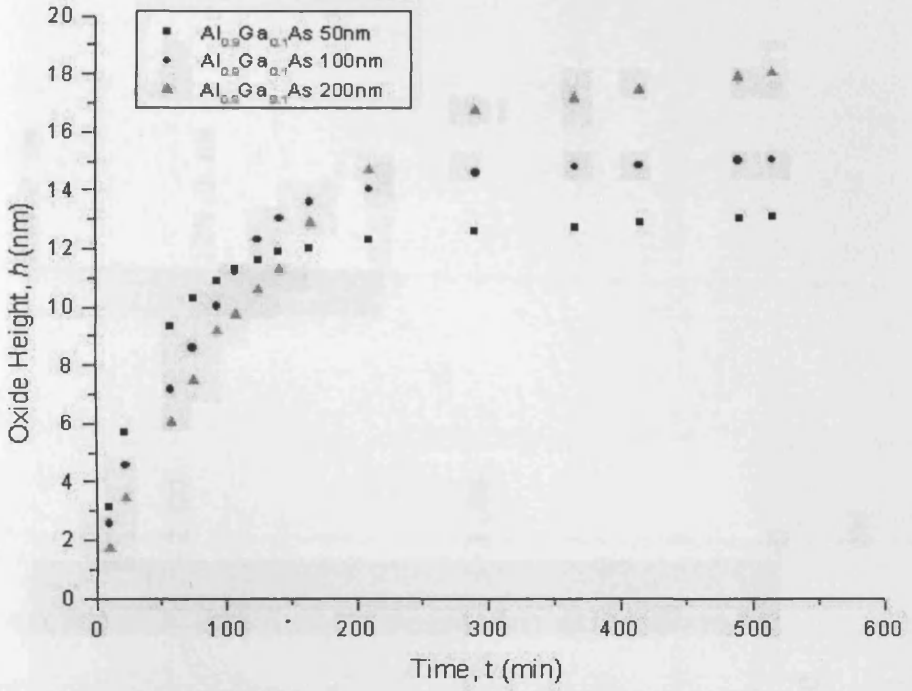

Fig 6.11 A plot of mean oxide height difference versus time for the $Al_{0.9}Ga_{0.1}As$ layers.

6.3.6 Measurements of AlAs Layers

Fig 6.12 shows a topographic image of the four AlAs layers of the sample 20 min after cleaving. Observed inhomogeneities in the oxide growth along the 200 and 400 nm thick layers became more pronounced with increasing oxidation. Such features were also observed in the $Al_xGa_{1-x}As$ layers of all devices (mainly VCSELs) where $x > 0.9$. It is difficult to speculate as to the cause of these inhomogeneities, but they could possibly be strain related.

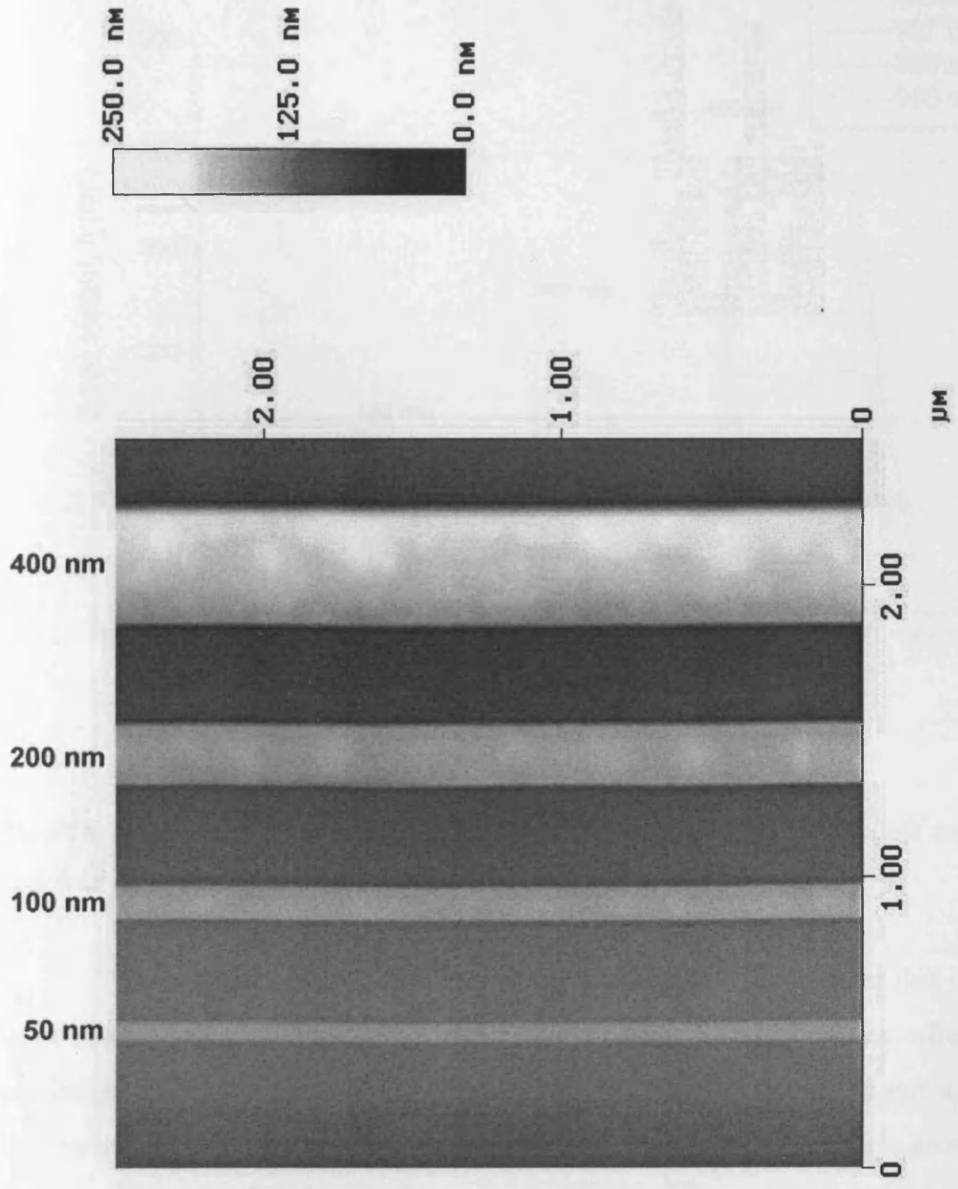

Fig 6.12 A topographic AFM image of the AlAs layers in the sample, after 20 min of oxidation. The surface of the sample is to the top of the page with the other $Al_xGa_{1-x}As$ layers to the bottom.

The evolution of the oxide profiles of these layers with time is illustrated in fig 6.13.

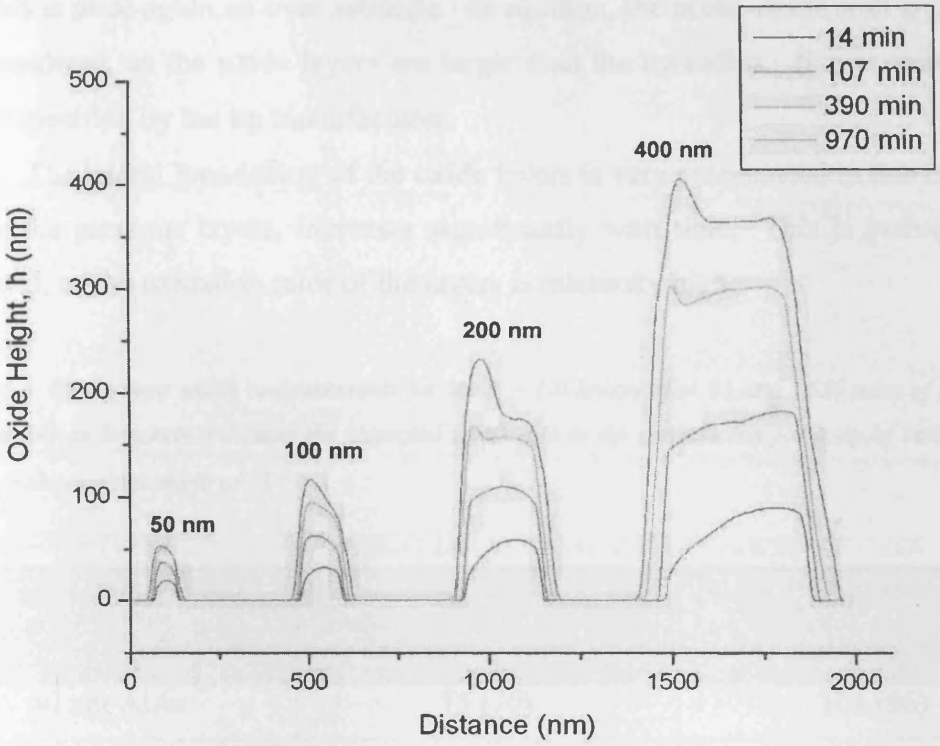

Fig 6.13 The evolution of the oxide profiles of the AlAs layers with time. The sample surface is 500 nm to the right of the 400 nm AlAs layer.

There were several interesting observations to be made from this data. Firstly, large bands of edge enhanced oxidation appear on the oxides after prolonged oxidation. These bands appeared throughout the sample and were not a result of the line averaging process 'smoothing over' inhomogeneities in the oxide growth. Indeed, they are clearly visible in AFM images that have not been presented here. This effect is difficult to explain, but may be related to dislocations in the layer or the close proximity of the surface. Secondly, although initially rounded, the oxide layers displayed an asymmetry that was not observed in the layers of lower aluminium concentration. In addition, the direction of this asymmetry changes as the growth proceeds, ruling out any tip convolution effects. Again, it is possible that the changes in oxide profile are related to strain or dislocations within the layer.

Oxide base width values after 15 and 1650 min of oxidation are presented in table 6.4. Once again, figures in brackets indicate the expected tip convolved value with a tip radius of 20 nm, although close inspection of the figures presented reveals that this is once again an over estimate. In addition, the probe-oxide wall angle must be considered, as the oxide layers are larger than the tip radius. It is assumed to be 11° as specified by the tip manufacturer.

The lateral broadening of the oxide layers is very pronounced in this case, and unlike the previous layers, increases significantly with time. This is perhaps to be expected, as the oxidation rates of the layers is relatively high.

Table 6.4 *Oxide base width measurements for the x = 1.0 layers after 15 and 1650 mins of oxidation. The number in brackets indicates the expected width due to tip convolution for a tip of radius 20 nm and a probe contact angle of 11°.*

Layer	Base Width (nm) ± 4 nm 15 mins after cleave	Base Width (nm) ± 4 nm 1650 mins after cleave
50 nm AlAs	75 (70)	101 (96)
100 nm AlAs	130 (142)	190 (169)
200 nm AlAs	234 (249)	360 (309)
400 nm AlAs	440 (459)	700 (591)

Fig 6.14 shows plots of the oxide height difference between oxidised AlAs and GaAs as a function of oxidation time for each of the four layers. The behaviour is as expected, with the wider layers oxidising the most quickly. The also data shows that extremely thick oxide layers grow on the material. Indeed, for the widest of the layers, the oxide height is comparable to the width of the layer itself, i.e. 400 nm.

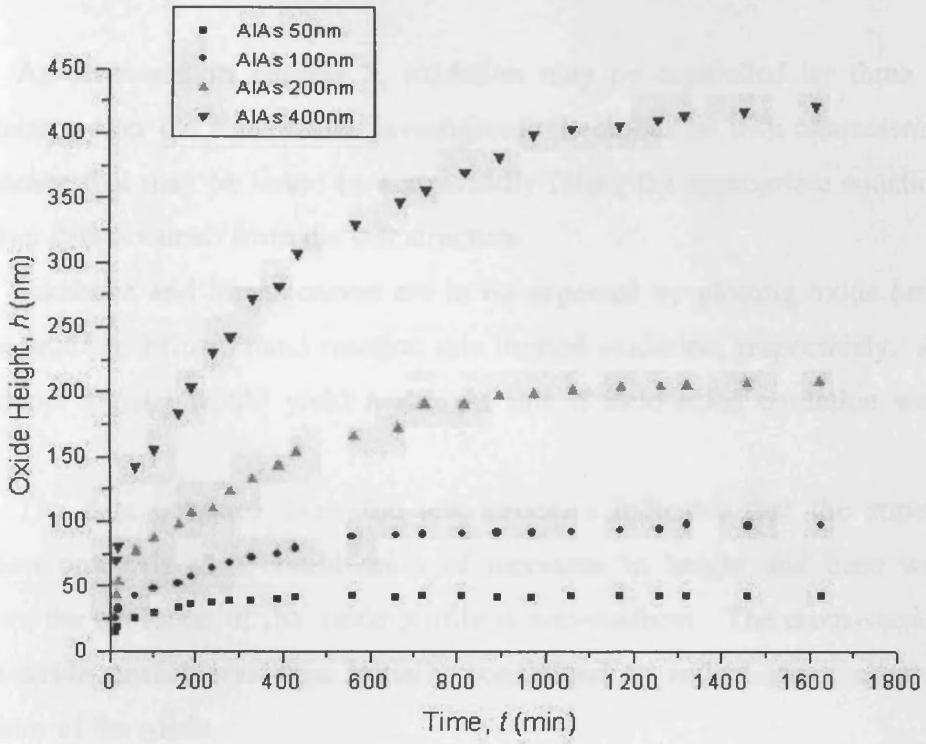

Fig 6.14 A plot of oxide height difference versus time for the AlAs layers.

6.4 Interpretation of Results

This section is devoted to the detailed discussion of the results obtained by the investigation of the test sample. Several interesting phenomena were observed that warrant explanation, and this is attempted where possible. The analysis begins by attempting to identify the oxidation regimes involved for each oxidising layer. The time evolution of the oxide profile is then outlined, with particular emphasis being placed on the observed crowning effect. The broadening of the oxide base width is then analysed, followed by a discussion of the relationship between oxidation rate and layer width.

6.4.1 Identification of Oxidation Mechanisms

As discussed in chapter 3, oxidation may be controlled by three possible mechanisms over the time scales investigated. Each has its own characteristic time dependency that may be found by successfully fitting the appropriate equation to the oxidation data obtained from the test structure.

Parabolic and linear curves are to be expected by plotting oxide height data against time for diffusion and reaction rate limited oxidation, respectively. A plot of $1/h$ against $\log_{10}(t)$ would yield a straight line if field-aided oxidation was taking place.

The data obtained from the test structure indicates that the super-surface oxidation proceeds as a combination of increases in height and base width. In addition, the evolution of the oxide profile is non-uniform. The cross-sectional area of the oxide profile was thus initially considered to reflect more accurately the behaviour of the oxide.

However, a comparison of the time evolution of both the oxide cross-sectional area and oxide height for several different layers showed the same dependence. It was thus concluded that there was no significant advantage in the rather laborious calculation of oxide cross-sectional area. Thus only the time dependence of oxide height data is considered during the identification of the different oxidation regimes.

As an example, the time evolution of oxide height and cross-sectional area for the 50 nm wide AlAs layer is presented in fig 6.15. The parabolic fits (red lines) in both graphs, suggest that the same oxidation behaviour is exhibited by both quantities. In this case, the parabolic dependence indicates diffusion rate limited oxidation is taking place.

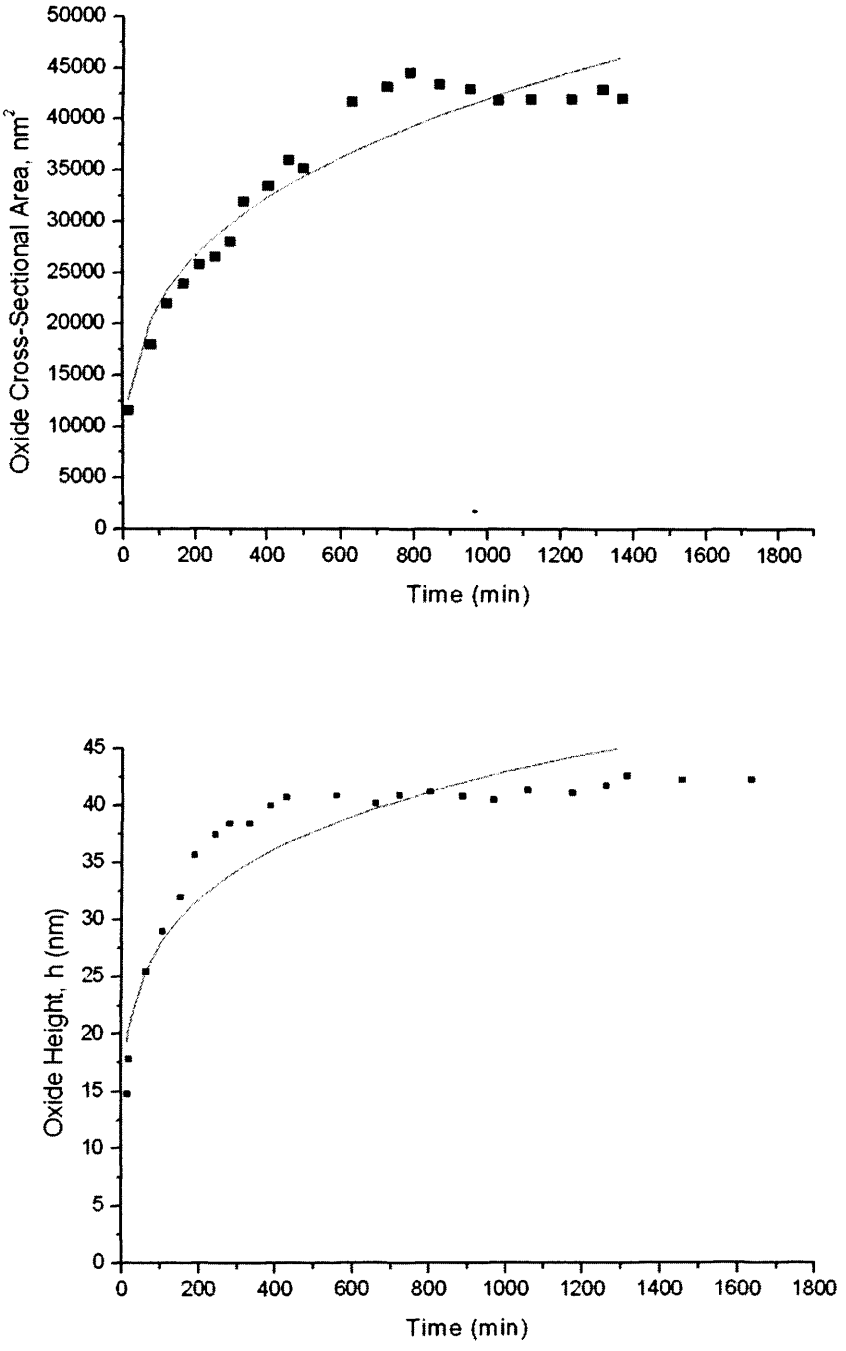

Fig 6.15 *The variation of oxide cross-sectional area (above) and height (below) with time for the 50 nm thick AlAs layer. The red lines in both graphs are parabolic fits, a comparison of which indicates similar behaviour for both quantities.*

6.4.1.1 Field-aided Oxidation

The criterion for field-aided oxidation was that a straight line could be fitted to a plot of $1/h$ against $\log_{10}(t/\tau)$ for the layer in question. This was found to be the case for the initial oxidation of all the layers within the range $x = 0.3$ - 0.8. Since the value of the time scaling constant τ was unknown, it was assumed to be unity in all cases. The effects of the crowned regions were not considered and are analysed separately later in this chapter. Height measurements were thus taken from the centre of the oxide growth for each layer, where enhanced oxidation had not taken place.

Fig 6.16 shows a plot of $1/h$ against $\log_{10}(t/\tau)$ 100 nm thick $Al_{0.3}Ga_{0.7}As$ layer, and is typical of all the layers whose data that exhibited this behaviour. A straight-line fit (red line) is possible for the data up to a certain time, after which the fit is no longer possible. At this point, there is a clear transition to another oxidation regime (discussed later).

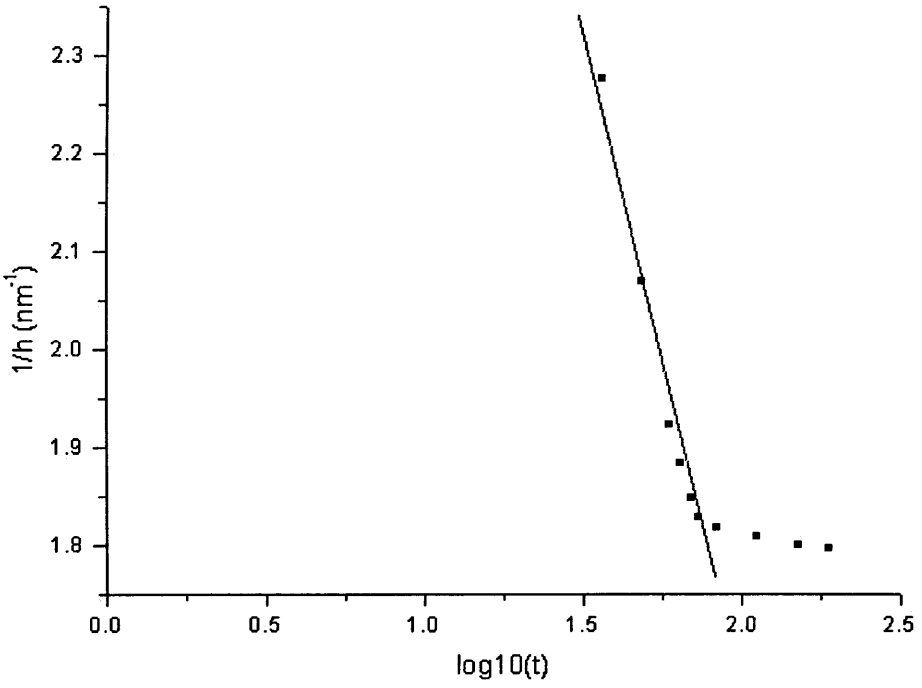

Fig 6.16 A plot of 1/h against $\log_{10}(t/\tau)$ for the 100 nm thick $Al_{0.3}Ga_{0.7}As$ layer.

Table 6.5 summarises the transition times from field-aided to the subsequent oxidation mode for the layers that exhibit this behaviour. As can be seen, the trend is for the transition times of the materials to increase with aluminium concentration and layer thickness. Composition differences appear to have a greater effect on the transition time than layer thickness. In addition, the differences in transition times between adjacent layers becomes greater as the aluminium content is increased. It should be noted however, that the effects of GaAs oxidation may not be insignificant for these layers, particularly those with a low aluminium concentration.

These results may be consistent with the findings of Reinhardt [1] (see section 3.2.3.2), where a range of transition times between 70 and 150 min were found for layers with mole fractions in the range x = 0.12 - 0.72 [2]. However, the previous work did not consider the effect of layer thickness on the transition times. Also, differences in layer thickness and composition make a direct comparison difficult.

Table 6.5 *The transition times between field-aided oxidation and subsequent rate limiting process for layers within the range $x = 0.3 - 0.7$.*

Layer	Transition Time (min) ± 10 min
50 nm $Al_{0.3}Ga_{0.7}As$	81
100 nm $Al_{0.3}Ga_{0.7}As$	93
50 nm $Al_{0.5}Ga_{0.5}As$	101
100 nm $Al_{0.5}Ga_{0.5}As$	105
50 nm $Al_{0.7}Ga_{0.3}As$	120
100 nm $Al_{0.7}Ga_{0.3}As$	132
50 nm $Al_{0.8}Ga_{0.2}As$	170
100 nm $Al_{0.8}Ga_{0.2}As$	184

It is also possible to extract a number of useful parameters relating to the oxidising material from this data. Comparing equation 3.1 with the equation for a straight line reveals that these parameters may be calculated from the gradient and y intercept of the $1/h$ against $\log_{10}(t/\tau)$ plots for this data. The y intercept yields a value of h_0, the oxide film thickness at which field-aided growth begins, while a value of h_1, a term proportional to the Mott Potential of the material, can be calculated from the

gradient. It should again be noted that the heights are measured with respect to GaAs, and that the time scaling constant τ was unknown, so accurate values are unlikely. However, it is possible to gain a correct order of magnitude for these quantities. The calculated values are presented in table 6.6.

Table 6.6 *h_1 and h_0 values for the layers within the range $x = 0.3 - 0.7$, as calculated from the plots of $1/h$ against $\log_{10}(t/\tau)$.*

Layer	h_0 (Å) (± 0.5Å)	h_1 (Å) ± 0.5Å
50 nm $Al_{0.3}Ga_{0.7}As$	1.9	6.1
100 nm $Al_{0.3}Ga_{0.7}As$	2.0	6.5
50 nm $Al_{0.5}Ga_{0.5}As$	2.5	7.2
100 nm $Al_{0.5}Ga_{0.5}As$	2.4	7.8
50 nm $Al_{0.7}Ga_{0.3}As$	2.8	8.8
100 nm $Al_{0.7}Ga_{0.3}As$	2.6	9.1
50 nm $Al_{0.8}Ga_{0.2}As$	3.0	12.8
100 nm $Al_{0.8}Ga_{0.2}As$	3.2	14.6

From this data, it can be seen that layer thickness has little or no effect on the initial oxide film thickness. The values obtained are perhaps a little larger than would be expected, but they are of the right order of magnitude, and they do reveal a trend for increasing film thickness with increasing aluminium concentration. As discussed in chapter 3, this is to be expected, as an increase in aluminium within the layer will increase its field strength.

While the values of h_1 are not directly useful, they allow the effect of layer width and material on the Mott Potential to be quantified. The data suggests that the Mott Potential increases with increasing layer width and aluminium concentration. This quantity is a property of the material, so while a dependence is expected for variations in aluminium mole fraction, it is not clear why the value should change with layer width. The most likely explanation relates to the mechanical stress placed on the constrained oxidation front, as outlined in section 3.3.1. It is possible that the stresses involved modify the electric field of the layer and hence the Mott Potential, creating a dependence on layer width.

As discussed earlier, straight lines could not be fitted to $1/h$ against $\log_{10}(t/\tau)$ plots for layers in the range $x = 0.9 - 1.0$. As a representative example, fig 6.17 shows this plot for the 400 nm wide AlAs layer. It was concluded that field-aided oxidation does not take place in these layers for the times at which the data was collected.

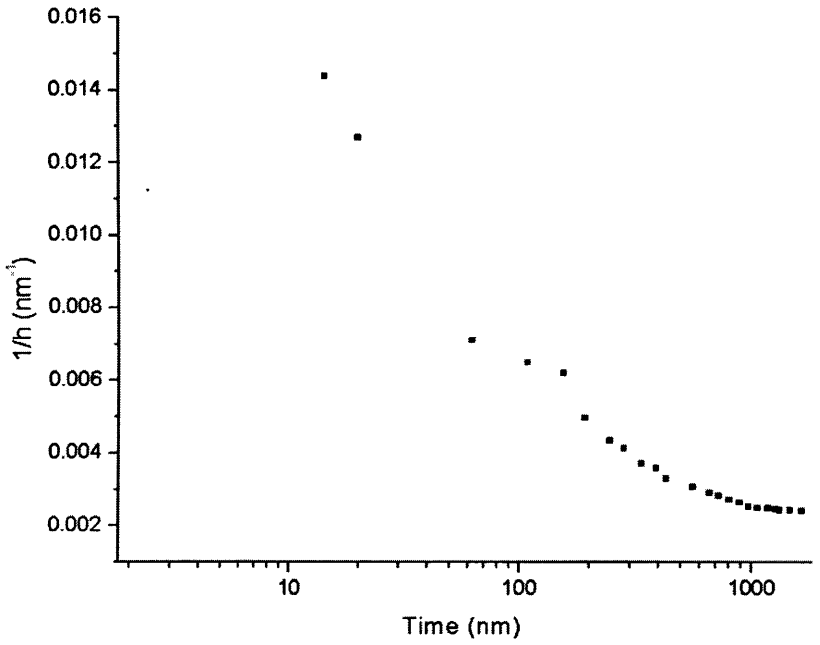

Fig 6.17 A plot of 1/h against $\log_{10}(t/\tau)$ for the 400 nm wide AlAs layer.

6.4.1.2 Reaction Rate Limited Oxidation

No evidence of reaction rate limited oxidation could be found for any of the layers of the test structure. Linear fits were attempted for all of the data, but failed in each case.

This is not surprising, however. As outlined in chapter 3, reaction rate limited oxidation has been shown to be favoured under conditions of elevated temperature and moisture. These experiments were carried out under ambient conditions, where reaction rate limited oxidation is not expected to be the dominant mode of oxidation. These results are therefore consistent with data from literature.

6.4.1.3 Diffusion Rate Limited Oxidation

Given the results outlined in the previous section, it is reasonable to assume that diffusion rate limited oxidation will be the dominant mode after the transition from field-aided oxidation. Indeed, this is found to be the case. Parabolic fits to all of the data showed that diffusion rate limited oxidation occurred for layers in the range $x = 0.3 - 0.8$ after field-aided oxidation has taken place, and that diffusion rate limited oxidation was the dominant oxidation mode for layers in the range $x = 0.9 - 1.0$ throughout the lifetime of the sample.

Fig 6.18 shows a parabolic fit to the data obtained from the 400 nm wide AlAs layer. The data does appear to deviate from ideal parabolic behaviour for large values of t, which may be due to the tip wearing down the oxide, as discussed in section 6.3.1. Since the oxidation rate had slowed down considerably by this time, tip-wearing effects would become more significant.

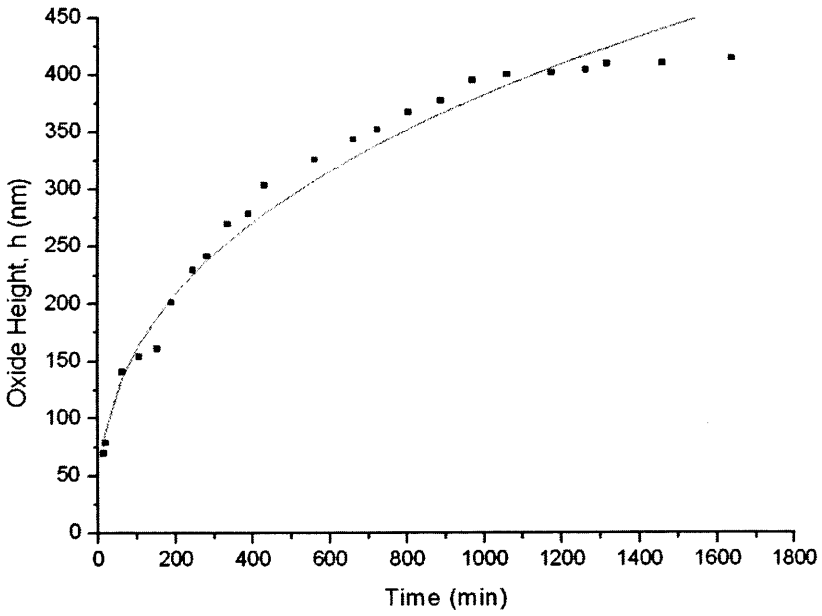

Fig 6.18 A plot of mean oxide height difference versus time for the 400 nm thick AlAs layer. The red line is a parabolic fit to the data.

As the AlAs layers became narrower, the parabolic fit became less convincing, as illustrated in fig 6.19 for the 50 nm wide layer. This can be attributed to the crowning effect outlined earlier. Evidence for this comes from the 100 and 200 nm thick $Al_{0.9}Ga_{0.1}As$ layers, where it was possible to separately measure the oxide height

at the edge and centre of the oxide layers, allowing a separate analysis of the crowning and bulk oxidation. Fig 6.20 plots these two sets of data for the 100 nm wide $Al_{0.9}Ga_{0.1}As$ layer. The green data points illustrate the time evolution of the central region of the oxide layer. The parabolic fit clearly shows that the central region does indeed fit the criterion for diffusion rate limited oxidation. The black data points illustrate the time evolution of the crowned edge regions. The oxidation mechanism here is clearly not diffusion rate limited. Comparison with the data presented in fig 6.18 confirms the assertion that the deviation from ideal parabolic behaviour is a result of the crowning effect. It should also be noted that the time at which the two sets of data in fig 6.20 began to converge coincided with the overlapping of the two crowned regions.

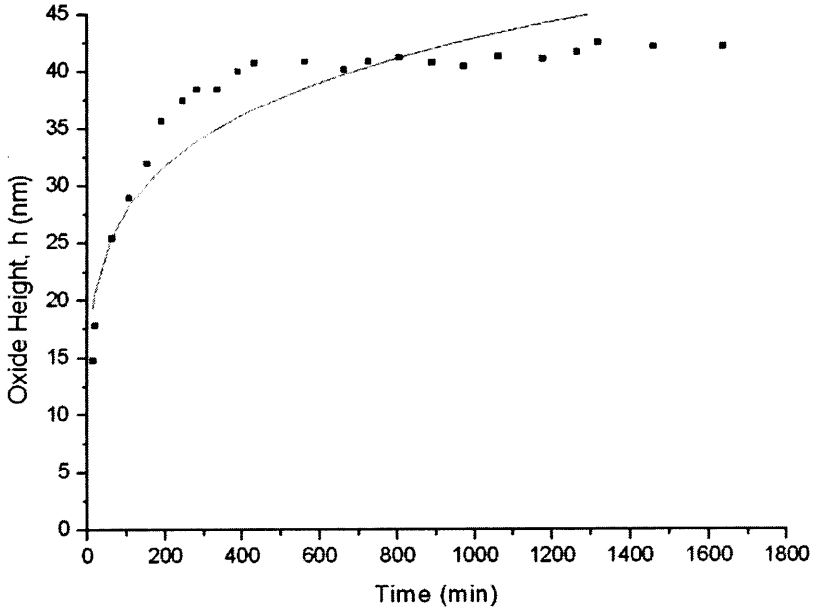

Fig 6.19 *A plot of mean oxide height difference versus time for the 50 nm thick AlAs layer. The red line is a parabolic fit to the data.*

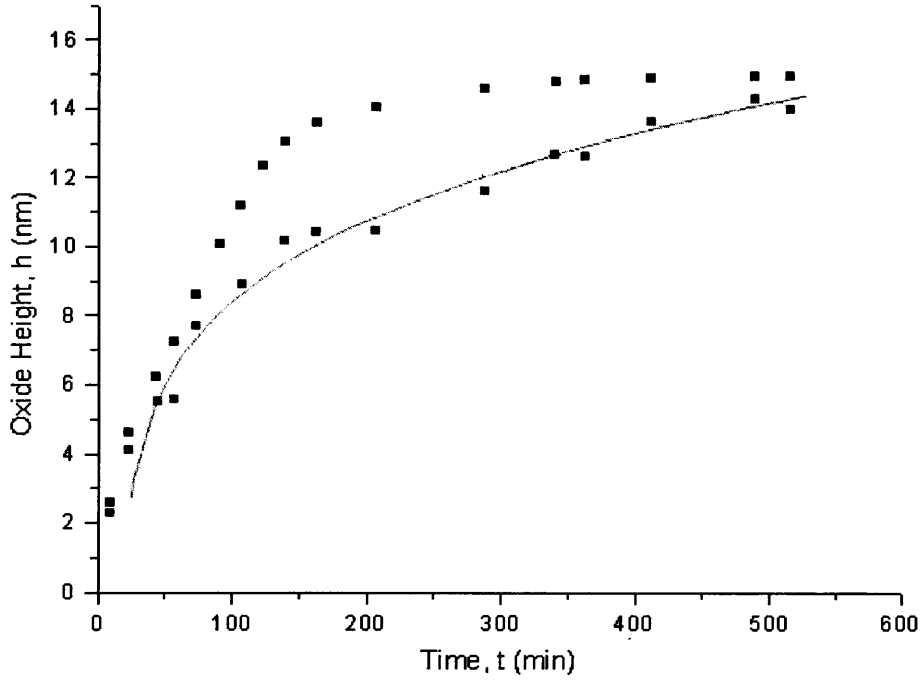

Fig 6.20 A plot of mean oxide height difference versus time for the 100 nm thick $Al_{0.9}Ga_{0.1}As$ layer. Black data points indicate the heights at the edge of the oxide layer. Green data points are those at the centre. The red line is a parabolic fit to the data, showing that the oxidation within the central region of the oxide layer is indeed diffusion rate limited.

As with the field-aided oxidation case, it is possible to obtain a useful parameter from the data obtained. From the parabolic plots presented here, the coefficient B (equation 3.12) as outlined in section 3.2.3.1 may be calculated. This coefficient is directly proportional to the diffusion coefficient D of the oxidising material. The values obtained for B will thus give an indication of the behaviour of the diffusion coefficient.

Values of B for the AlAs layers are presented in table 6.6. Values obtained for the other layers within the test structure were thought to be unreliable due to the presence of crowning effects and the transition from field-aided oxidation within the data.

Table 6.6 B values for the four AlAs layers of the test structure, as calculated from the parabolic fits to the oxide height v's time data.

Layer	B (nm min$^{-½}$) ± 16.4 × 10^{-5} nm min$^{-½}$
400 nm AlAs	16.4×10^{-5}
200 nm AlAs	8.7×10^{-5}
100 nm AlAs	4.2×10^{-5}
50 nm AlAs	2.1×10^{-5}

The values of B are larger for the thicker layers, doubling as the layer thickness doubles. This implies that the diffusion coefficient and hence the oxidation rate will also double as the layer with is doubled. Evidence for this is presented in fig 6.12, where it can be clearly observed that the oxide height at a particular time doubles as the layer width is doubled.

There is no reason to expect a dependence of diffusion coefficient on layer width, as was the case earlier for the Mott Potential. The reason is again possibly concurrent with the theories outlined in section 3.3.1, with the diffusion coefficient being dependent on the mechanical stress within the constrained oxidising layer.

6.4.2 Crowning Effects and the Evolution of the Oxide Profile

The evolution of the oxide profile has already been partly explained during the identification of the oxidation mechanisms as presented above. This section aims to bring together this knowledge and present a model for the time evolution of the oxide profiles.

It has been found that the evolution of the profiles is governed by more than the transitions between successive oxidation mechanisms. The crowning effect observed for the majority of the layers appears to be the dominant factor in determining the shape of the profile. This crowning has been observed in previous literature [2]. However, in that instance it was not commented upon, and its significance was clearly not appreciated. The crowned profiles indicate that some form of *interface enhanced oxidation* is taking place. Data from the test structure

suggests that it can not be attributed to any of the three different oxidation mechanisms considered previously.

It is beyond the scope of this thesis to present a theoretical model for the interface enhanced oxidation. However, phenomenologically it is dependent on both layer width and composition, a property it shares with other forms of oxidation. An increase in aluminium concentration increased the interface enhanced oxidation rate. However, unlike the other forms of oxidation, it is inversely proportional to the layer width. Evidence also indicates that although the oxidation mechanism involved contributes the majority of the initial oxide height increase, the oxidation slows down considerably as the two crowned regions overlap and merge. This allows the oxidation to become diffusion rate limited.

When compared with the oxidation front model outlined in section 3.3.1.1, it appears the two are not compatible. The proposed shape of the advancing oxidation front within the layer suggests that there should in fact be less oxide growth at the edges of the layer. There are however a number of interface effects that have not been considered in the model, which may be important contributory factors to interface enhanced growth. The electric field is discontinuous at the interface and there are also chemical differences to consider. These could both increase the mobility of the oxidation species. There may also be defects and strain within the layers, particularly those with higher aluminium concentrations.

The presence of interface enhanced oxidation can be used to explain the behaviour of the $Al_{0.9}Ga_{0.1}As$ layers. As outlined earlier in this chapter, the narrower layers of this material initially have the thicker oxide growth. As the oxidation progresses however, the oxide growth on the wider layers eventually becomes thicker. The inverse proportionality to the layer thickness of the interface enhanced oxidation encourages a thicker oxide layer to build up on the narrower layers. However, as the crowned regions merge (as is the case of the 50 nm layer in fig 6.10), this form of oxidation slows down, allowing diffusion rate limited oxidation to become the dominant mode. In addition, the merging of the two crowned regions takes less time for the narrower layers. The dependence of diffusion rate limited oxidation on layer width eventually encourages a thicker oxide to grow on the wider layers.

Fig 6.21 illustrates the evolution of the oxide profile on an oxidising layer (grey) with time. At time (a), the interface enhanced growth (black) produces the typical crowned profile. This region of oxidation is separated by a central area where

either field-aided or diffusion rate limited oxide growth (red) takes place. Over time (b-c) each oxidation mechanism contributed to the evolution of the profile until finally the interface enhanced oxide regions overlap (d), producing a round-topped profile. At this point, the slowing of the interface enhanced oxidation rate allows the other oxidation mechanism to dominate.

In the model below, it is assumed that the field-aided and diffusion limited oxidation mechanisms produce a flat-topped profile. Results presented earlier show that this is the case for field-aided growth. The presence of interface induced oxide growth makes it impossible to determine if this is true diffusion rate limited regime. However, it has been assumed in order to simplify the model.

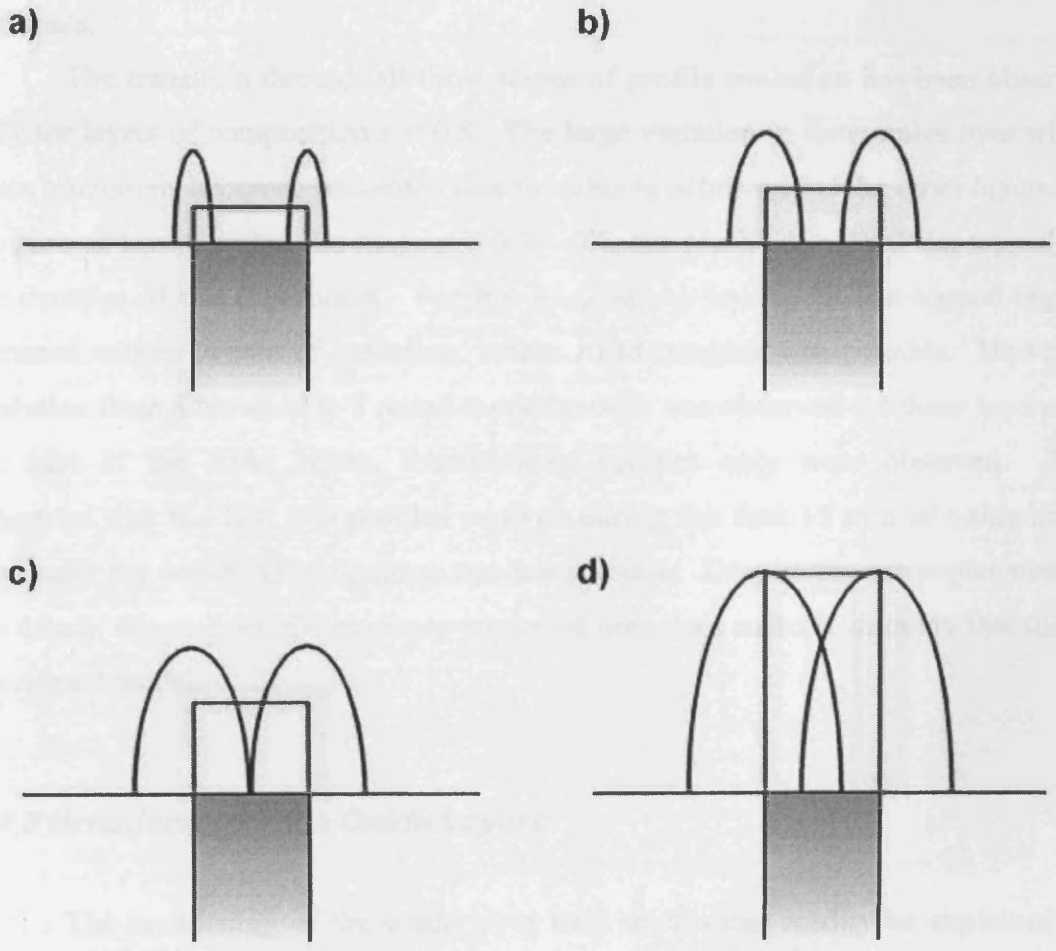

Fig 6.21 *The evolution of the oxide growth on an AlGaAs layer (grey). The profile of interface enhanced oxidation is shown by the elliptical solid black lines. The profile of the oxide growth due to field-aided or diffusion rate limited oxidation is shown in red. As the oxidation continues through various times* ***(a-d)****, the crowned regions eventually merge, producing a round-topped profile* ***(d)****.*

It is thus proposed that the evolution of the oxide profile involves a three-step process. In the initial stages of oxidation, the oxide profile is flat-topped and the oxidation proceeds via the field aided mode. At some point, depending on both the material and layer width, a transition to diffusion limited oxidation occurs. As these mechanisms drive the oxidation of the bulk of the layer, an independent process fuels an enhanced oxidation of interface regions. This mechanism is unrelated to the two that govern the oxidation of the bulk of the layer, and initially occurs at a much faster rate to the other two mechanisms, producing a crowned profile. As the oxidation continues, the two peaks of the crowed regions overlap producing a round-topped oxide profile. At this point, the interface enhanced oxidation slows, allowing the oxidation mechanism of the bulk of the material (by this time diffusion limited) to dominate.

The transition through all three stages of profile evolution has been observed only for layers of composition $x = 0.8$. The large variation in time scales over which these transitions occurred prevented this from being achieved for the other layers. In the case of layers within the range $x = 0.3 - 0.7$, the profile remained flat topped for the duration of the experiment. For the $Al_{0.9}Ga_{0.1}As$ layers, the flat-topped regime occurred within 15 min of oxidation, before AFM imaging was possible. However, evolution from a crowned to a round-topped profile was observed for these layers. In the case of the AlAs layers, round-topped profiles only were observed. This suggested that the first two profiles evolved during the first 15 min of oxidation, a time scale for which AFM imaging was not possible. Despite the incompleteness of the data in this respect, the evidence presented here does indicate strongly that this is the correct model.

6.4.3 Broadening of the Oxide Layers

The broadening of the oxide layer base widths can readily be explained by interface enhanced oxidation. The model outlined in fig 6.21 clearly illustrates how the broadening occurs. As the crowned regions grow, the oxide spreads outwards. Oxide layers measured that do not show signs of interface enhanced growth exhibit hardly any increase in lateral dimensions. It is not possible to state whether there is a broadening of the oxide without the presence of interface enhanced growth however.

The ambiguities in oxide base width measurements due to tip convolution as discussed in section 6.3.4 make it impossible to determine. Field-aided and diffusion limited oxidation therefore may or may not produce any increase in the lateral dimensions of the oxide layer.

For layers in the range $x = 0.9 - 1.0$, the lateral broadening of the oxide layer due to interface enhanced oxidation is very pronounced indeed, and increases with layer width and aluminium mole fraction. Unfortunately, as discussed earlier, no trends related to the interface enhanced oxidation could be discerned from the data. It is therefore not possible to comment on the laws that govern this oxidation mechanism.

6.4.4 Dependence of Oxidation Rate on Layer Thickness

From the data collected, it is possible to comment upon the relationship between oxidation rate and layer thickness. For the field-aided regime, the data suggests a complex relationship. However, the uncertainties associated with interface enhanced oxidation inhibit any confident statements to be made as to the nature of this relationship.

For the AlAs layers however, this is not an issue, as diffusion rate limited oxidation dominates. It has already been established (section 6.4.3.1) that a simple relationship exists between the two quantities. As the layer thickness doubles, so too does the oxidation rate.

It is also possible also to comment on the effect the layer thickness has on oxidation mechanism. It has been shown that the layer thickness governs the transition time between successive oxidation mechanisms, not the oxidation mechanism itself. It should be noted however, that this is only true for the ambient conditions under which these experiments were conducted. As discussed in chapter 3, it is widely reported in the literature that the rate limiting process is highly dependent on the layer width at elevated temperatures.

The layer thickness has also been shown to govern the oxidation rate for each of the three processes considered. The model put forward explaining the effect of layer thickness on oxidation rate has already been outlined in section 3.3.1. It considers the effect of interfacial energies and the mechanical stress at the curved

oxidation front. However, the effects are only proposed to modify the energies of the oxidation reaction, and thus the reaction rate limiting process.

Despite the fact the Mott Potential and the diffusion coefficient are material properties, it possible that the same mechanical stresses are responsible for modifying their values. However, further speculation will not be attempted here.

6.5 Composition Determination and Interface Positioning

After the data analysis had been completed, it was important to establish how it could be used practically to aid with the metrology of real device structures. Data is presented over 1700 min, but in an industrial environment it is not viable to collect data for such a long period of time. It was therefore decided that the data be collated as to allow the determination of layer thickness and composition after one hour of oxidation.

It has already been shown that the thickness of layers can be determined to within 3 nm by measuring the sample and taking the differential of the oxide profiles. Composition determination may be achieved by comparing measured maximum oxide heights after one hour with the data presented in fig 6.22. This graph plots oxide height as a function of layer thickness for the available compositions of $Al_xGa_{1-x}As$. The height is plotted on a logarithmic scale so that all the data fits on a single graph. Data for $Al_{0.1}Ga_{0.9}As$ was not included here due to the fact no discernible oxide difference was observed at this time.

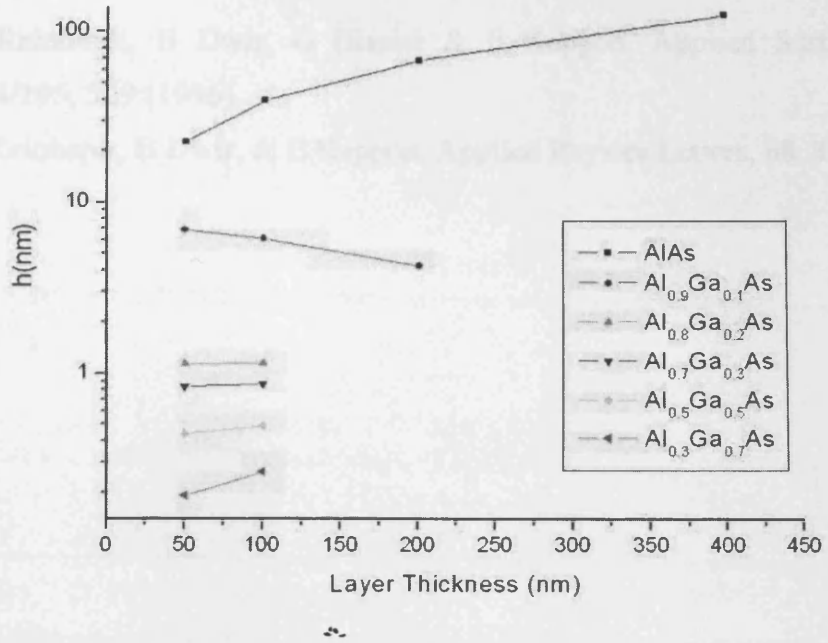

Fig 6.22 *The evolution of oxide height as a function of layer thickness for various compositions of $Al_xGa_{1-x}As$ as determined from the test structure, after one hour of oxidation. Heights are measured from the highest point of oxidation to the level of the surrounding GaAs.*

Given the layer thickness can be determined to within 3 nm, aluminium concentrations x may be determined to within 0.05 for the mole factions $x < 0.8$. For materials where $x > 0.8$, where relative changes in the oxidation rate are greater, x may be determined to within 0.02. For epitaxial growth, the composition of AlAs is almost certainly correct. This is a result of the fact that GaAs is not present when the material is grown.

Clearly, the data presented here is far from complete. However, a certain amount of extrapolation may be applied to the graph above, leading to results that are well within errors resulting from the original measurements. As an example, a 75 nm layer having undergone 6.5 nm of oxidation after one hour may been seen to be of composition $x = (0.90 \pm 0.02)$.

Finally, it should be remembered that the relationship between oxide height and layer thickness in this graph are only valid for one hour of oxidation. Different times require a different graph.

6.6 References

[1] F Reinhardt, B Dwir, G Biasiol & E Kappon. Applied Surface Science, **104/105**, 529 (1996).

[2] F Reinhardt, B Dwir, & E Kappon. Applied Physics Letters, **68**, 3168 (1996).

Chapter 7

Studies of Device Structures

7.1 Introduction

The aim of this chapter is to illustrate how effectively AFM may be used to characterise real device structures, and identify typical problems that may occur during epitaxial growth and device fabrication. To this end, several contrasting studies are presented here. The majority of samples studied were $Al_xGa_{1-x}As/GaAs$ based so that direct comparisons with the test sample outlined in chapter 6 may be made. However, two of the samples studied were phosphide based. The lack of comparative data from a test sample such as that of the $Al_xGa_{1-x}As/GaAs$ materials meant that no determination of composition or layer thickness could be made. However, phosphide based materials are in widespread use in III-V devices, and results presented indicate that this technique can be applied successfully to them. Where possible, the results obtained are compared to those derived by other characterisation techniques, and the relative merits of AFM commented upon.

Section 7.2 begins with several studies of $Al_xGa_{1-x}As/GaAs$ based device structures, beginning with two typical VCSELs in section 7.2.1. A third VCSEL is studied in section 7.2.2. This one has been processed into a device, and in particular has undergone lateral steam oxidation. The determination of the position and thickness of incorrectly grown layers in a laser diode is outlined in section 7.2.3, while the treatment of $Al_xGa_{1-x}As/GaAs$ based devices concludes with an examination of topographic differences between substrates and buffer layers.

A phosphide based complex multiple quantum barrier test structure is studied in section 7.3.1. Measurements of strain relaxed layers in a sample are made in section 7.3.2. Finally, a chapter summary is presented in section 7.4.

7.2 Investigation of $Al_xGa_{1-x}As$/GaAs based Devices

7.2.1 Investigation of VCSEL Structures

7.2.1.1 Background

VCSELs are devices where the layer thicknesses need to be very accurately controlled in order for the device to function. Metrology was performed on two such device structures grown by MOVPE at International Quantum Epitaxy (IQE) Ltd. It had been previously established that VCSELs fabricated from the same wafers operated correctly, indicating that layer thicknesses and compositions were accurate.

The specification for the samples is illustrated in tables 7.1 and 7.2. Errors in thickness and composition are those specified by IQE.

Table 7.1 *Specification for VCSEL sample 1-3623A. The surface of the sample is at the top of the table with the substrate at the bottom. Quantum well layers are marked *.*

Layer	Material	Composition (x) ± 0.02	Thickness (μm) ± 10 %
13	GaAs		0.0100
12	$Al_xGa_{1-x}As$	0.15	0.1823
11×30	$Al_xGa_{1-x}As$	0.15	0.0608
10×30	$Al_xGa_{1-x}As$	0.80	0.0687
9	$Al_xGa_{1-x}As$	0.50	0.0412
8	$Al_xGa_{1-x}As$	0.30	0.0500
7×3*	GaAs		0.0070
6×3	$Al_xGa_{1-x}As$	0.30	0.0110
5	$Al_xGa_{1-x}As$	0.30	0.0500
4	$Al_xGa_{1-x}As$	0.50	0.0412
3×39	$Al_xGa_{1-x}As$	0.80	0.0687
2×39	$Al_xGa_{1-x}As$	0.15	0.0608
1	GaAs		0.5000
S/S	GaAs		650.00

*Table 7.2 Specification for VCSEL sample 1-3280 D. The surface of the sample is at the top of the table with the substrate at the bottom. Quantum well layers are marked *.*

Layer	Material	Composition (x) ± 0.02	Thickness (μm) ±10%
21	GaAs		0.0050
20	$Al_xGa_{1-x}As$	0.12	0.0440
19	$Al_xGa_{1-x}As$	1.00>0.12	0.0200
18×19	AlAs		0.0509
17×19	$Al_xGa_{1-x}As$	0.12>1.00	0.0200
16×19	$Al_xGa_{1-x}As$	0.12	0.0389
15×19	$Al_xGa_{1-x}As$	1.00>0.12	0.0200
14	AlAs		0.0509
13	$Al_xGa_{1-x}As$	0.30<0.60	0.0870
12	$Al_xGa_{1-x}As$	0.30	0.0200
11×2	$Al_xGa_{1-x}As$	0.30	0.0080
10×3*	GaAs		0.0060
9	$Al_xGa_{1-x}As$	0.30	0.0200
8	$Al_xGa_{1-x}As$	0.60>0.30	0.0870
7	AlAs		0.0509
6	$Al_xGa_{1-x}As$	0.12<1.00	0.0200
5×30	$Al_xGa_{1-x}As$	0.12	0.0389
4×30	$Al_xGa_{1-x}As$	1.00>0.12	0.0200
3×30	AlAs		0.0509
2×30	$Al_xGa_{1-x}As$	0.12>1.00	0.0200
1	GaAs		0.5000
S/S	GaAs		650.00

The first VCSEL, sample 1-3623A, contained $Al_{0.80}Ga_{0.20}As/Al_{0.15}Ga_{0.85}As$ layers in the Bragg mirror stacks, of thicknesses 68.7/60.8 nm. No grades were present between the individual Bragg layers. The mirrors contained 39 repeats in the *n*-stack and 30 in the *p*-stack. The active region consisted of three 7 nm GaAs quantum wells, separated by 11 nm wide $Al_{0.3}Ga_{0.7}As$ spacers.

The second VCSEL, sample 1-3280D, contained $AlAs/Al_{0.12}Ga_{0.88}As$ layers in the Bragg mirror stacks, of thicknesses 50.9/38.9 nm. Graded regions of thickness 20 nm were present between the individual Bragg layers. The mirrors contained 30 repeats in the *n*-stack and 19 in the *p*-stack. The active region consisted of three 6 nm GaAs quantum wells, separated by 8 nm $Al_{0.3}Ga_{0.7}As$ spacers.

7.2.1.2 AFM Analysis of Sample 1-3623A

Measurements of the $Al_{0.8}Ga_{0.2}As$ Bragg layers after one hour reveal oxide heights of (1.30 ± 0.05) nm. The source of error here derives from the range of heights measured for 10 different layers. In this case, measurements were made with respect to $Al_{0.15}Ga_{0.85}As$, which may be assumed to have oxidised approximately to the same extent as GaAs at this time. This assumption was supported by the observation that no interface could be detected between GaAs and $Al_{0.15}Ga_{0.85}As$ layers at this time. Comparing the oxide height and layer width with results from the test sample in section 6.5, fig 6.19 leads to a layer composition of x = (0.80 ± 0.05), agreeing within errors with the specified value (0.80 ± 0.02).

The width of the layers was determined directly by calculating the differential of the oxide profiles. The width of the $Al_{0.8}Ga_{0.2}As$ oxide layers was measured to be (68 ± 3) nm. This is in good agreement within errors with the IQE specification of (68.7 ± 7.0) nm. The width of the $Al_{0.15}Ga_{0.85}As$ was measured to be (62 ± 3) nm, again in agreement with the specified value (60.8 ± 6.1) nm within growth tolerances. The error in both these cases was derived from the results obtained from 10 different layers.

It should be noted that the differential peaks that allow the determination of the layer thicknesses remain prominent for a considerable amount of time after cleaving. However, if the broadening from adjacent high mole fraction layers is allowed to merge, as has been observed, the differential peaks become small and wide. In this case, the determination of the individual layer thicknesses to within a tolerable error margin becomes impossible.

In addition to the measurement of the DBR layers, resolution of the individual quantum wells within the active region was possible. Fig 7.1 shows a topographic image of the active region of sample 1-3623A 2 hours after cleaving. The *n* and *p* mirror stacks are visible to the top and bottom of the image, with cleave damage evident in the upper *p*-mirror. The three quantum wells are visible as dark narrow bands in the centre of the image. The measured width was (6.4 ± 0.6) nm, the errors being derived from the spread in measurements from the three wells.

This result agrees within errors with the specified thickness of (7.1 ± 0.7) nm. The profile differential in this case revealed nothing, and these measurements are

therefore subject to errors introduced by the broadening of the adjacent $Al_{0.3}Ga_{0.7}As$. It was not possible to predict the extent of broadening expected for the layers between the quantum wells, as no data exists for layers that thin (11 nm). However, comparison between the profile differential and base width measurements for the 50 nm thick $Al_{0.3}Ga_{0.7}As$ spacers between the quantum wells and the DBR layers show the oxide has broadened by (5.0 ± 0.6) nm. This broadening could explain the low measured value of the quantum well thickness and the large (10 %) error.

The observed oxide step height between the wells and the surrounding material was 2 Å. Using equation 5.2, a probe tip radius of approximately 3.5 nm is required to reach the bottom of the quantum well. Since the tip radius is almost certainly larger than this value, it can be concluded that the measured depth is convolved, and is less than the actual value.

This observed contrast disappeared after a day of oxidisation, presumably due to the deleterious effect of the oxide broadening of adjacent layers as was expected. It is important therefore, that metrology on quantum well layers be conducted within a few hours after cleaving to avoid this broadening effect.

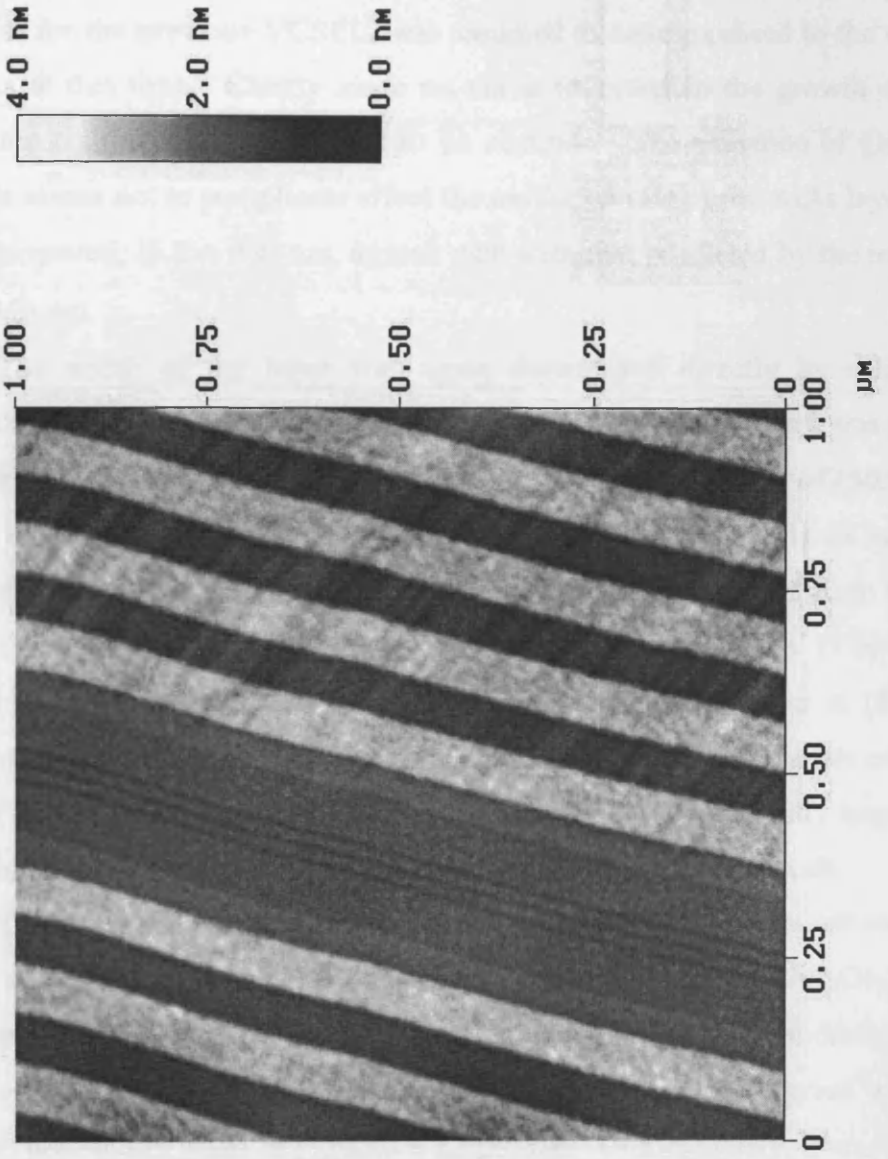

Fig 7.1 *A topographic AFM image of the area surrounding the active region of VCSEL 1-3623A after 2 hours of oxidation. The three quantum wells appear as thin dark bands at the centre of the image. The thick alternating light and dark bands are the high and low mole fraction layers of the DBR respectively. Note the appearance of cleave damage on the upper p stack.*

7.2.1.3 AFM Analysis of Sample 1-3280D

Measurements of the AlAs Bragg layers after one hour reveal oxide heights of (16 ± 1) nm. The source of error here is the range of heights measured for 10 different layers. In this case, measurements were made with respect to $Al_{0.12}Ga_{0.88}As$, which, as for the previous VCSEL, was assumed to have oxidised to the same extent as GaAs at this time. Clearly since no Ga is involved in the growth of the AlAs layers, the composition is expected to be accurate. The presence of grades in this structure seems not to complicate effect the oxidation rate of the AlAs layer, since the height measured, (6.2 ± 0.4) nm, agreed well with that predicted by the test structure, (6.0 ± 0.2) nm.

The width of the layer was again determined directly by calculating the differential of the oxide profiles. The width of the AlAs oxide layers was measured to be (58 ± 7) nm. This value is larger than the specified thickness of (50.9 ± 5.1) nm quoted by IQE. However, it is likely that the larger value is as result of the surrounding graded regions. Further evidence for this is obtained from the fact that the base width of the differential interface peak is larger than usual (7 nm) leading to the higher error quoted above. The width of each graded layer is (20 ± 2) nm. Assuming the grade is linear, then significant changes in the oxidation rate are likely to occur 3 nm either side of the AlAs layer. This explains the slightly larger measured layer width, and the larger base width to the differential interface peak.

The presence of grades in this sample meant that it was not reasonable to attempt direct measurement of the thickness of the intermediate $Al_{0.12}Ga_{0.88}As$ layers. However, a value of (134 ± 6) nm was obtained for the periodicity, the errors reflecting the spread in periodicity for individual layers. This agrees well with the specified periodicity value of (129.80 ± 13.00) nm. Unfortunately, it was not possible to determine the layer thickness beyond this.

As in the case of VCSEL 1-3623A, resolution of the individual quantum wells within the active region was possible. The measured width was (5.2 ± 0.5) nm, the errors being derived from the spread in measurements from the three wells. This result is smaller than the specified thickness of 6 nm, but is again likely to have been effected by the broadening of the adjacent $Al_{0.3}Ga_{0.7}As$.

The observed oxide step height between the wells and the surrounding material after an hour of oxidation was 0.6 Å, much less than that observed with the same materials for VCSEL 1-3623A. However, the narrower wells suggest that even less of the tip would be able to penetrate to the bottom, and it is concluded that the measured depth is a result of tip convolution.

7.2.2 Investigation of a Steam Oxidised VCSEL based Mesa Device

7.2.2.1 Background

This particular investigation was carried out in collaboration with the Optoelectronics Group at University of Wales Cardiff. Their interest lay in the use of lateral steam oxidation to produce current apertures in VCSEL devices.

The VCSEL structure, grown by MOVPE at IQE and designated QT1500, is presented in table 7.3. The lower *n*-Bragg stack consisted of $Al_{0.91}Ga_{0.09}As/Al_{0.15}Ga_{0.85}As$ layers. The top layers of the upper *p*-stack were of the same composition, but contained five layers that consisted of $AlAs/Al_{0.15}Ga_{0.85}As$. It was the steam oxidation of the AlAs layers that lead to the production of the current aperture, as outlined in section 2.9.3. The particular device in this investigation had a 50 μm wide mesa, produced by etching the sample down to the active region. Oxidation was carried out in a steam environment at 450°C for 20 min. Oxidation took place from the sides of the mesa inwards. At the temperature selected, it was thought that an oxide depth rate of 1 μms^{-1} would be achieved, resulting in the production of a current aperture whose target diameter was 20 μm.

It was possible to determine the extent of oxidation by cross-sectional AFM using topographic differences between layers that had undergone steam oxidation and those that had not. After cleaving, the layers that had not been steam oxidised would undergo subsequent oxidation. However, those having undergone steam oxidation already would be unable to oxidise further, resulting in topographical differences. In addition, the cleaved surface of non-single crystal oxide would be expected to be different to that of cleaved epitaxial material

*Table 7.3 Specification for VCSEL sample QT1500. The surface of the sample is at the top of the table with the substrate at the bottom. Quantum well layers are marked *.*

Layer	Material	Composition (x)	Thickness (Å)
30	GaAs	-	100
29	$Al_xGa_{1-x}As$	0.15	430
28×10	$Al_xGa_{1-x}As$	0.48	150
27×10	$Al_xGa_{1-x}As$	0.91	484
26×10	$Al_xGa_{1-x}As$	0.48	150
25×10	$Al_xGa_{1-x}As$	0.15	430
24	$Al_xGa_{1-x}As$	0.48	150
23×5	AlAs	-	491
22×5	$Al_xGa_{1-x}As$	0.48	150
21×5	$Al_xGa_{1-x}As$	0.15	430
20×5	$Al_xGa_{1-x}As$	0.48	150
19	$Al_xGa_{1-x}As$	0.91	484
18	$Al_xGa_{1-x}As$	0.48	150
17	$Al_xGa_{1-x}As$	0.15	430
16	$Al_xGa_{1-x}As$	0.48	150
15	$Al_xGa_{1-x}As$	0.91	643
14	$Al_xGa_{1-x}As$	0.28	900
13	GaAs*	-	94
12	$Al_xGa_{1-x}As$	0.28	98
11	GaAs*	-	94
10	$Al_xGa_{1-x}As$	0.28	98
9	GaAs*	-	94
8	$Al_xGa_{1-x}As$	0.28	900
7	$Al_xGa_{1-x}As$	0.91	643
6×20	$Al_xGa_{1-x}As$	0.48	150
5×20	$Al_xGa_{1-x}As$	0.15	430
4×20	$Al_xGa_{1-x}As$	0.48	150
3×20	$Al_xGa_{1-x}As$	0.91	484
2	$Al_xGa_{1-x}As$	0.48	150
1	$Al_xGa_{1-x}As$	0.15	430

Prior to investigation by AFM, scanning electron microscopy (SEM) was conducted on a different 50 μm wide mesa device, fabricated from the same wafer, and subjected to the same steam oxidation conditions. An SEM image of the cleaved cross-section of the edge of the mesa device is shown in fig 7.2. The resolution of this image was close to the limit achievable with the technique.

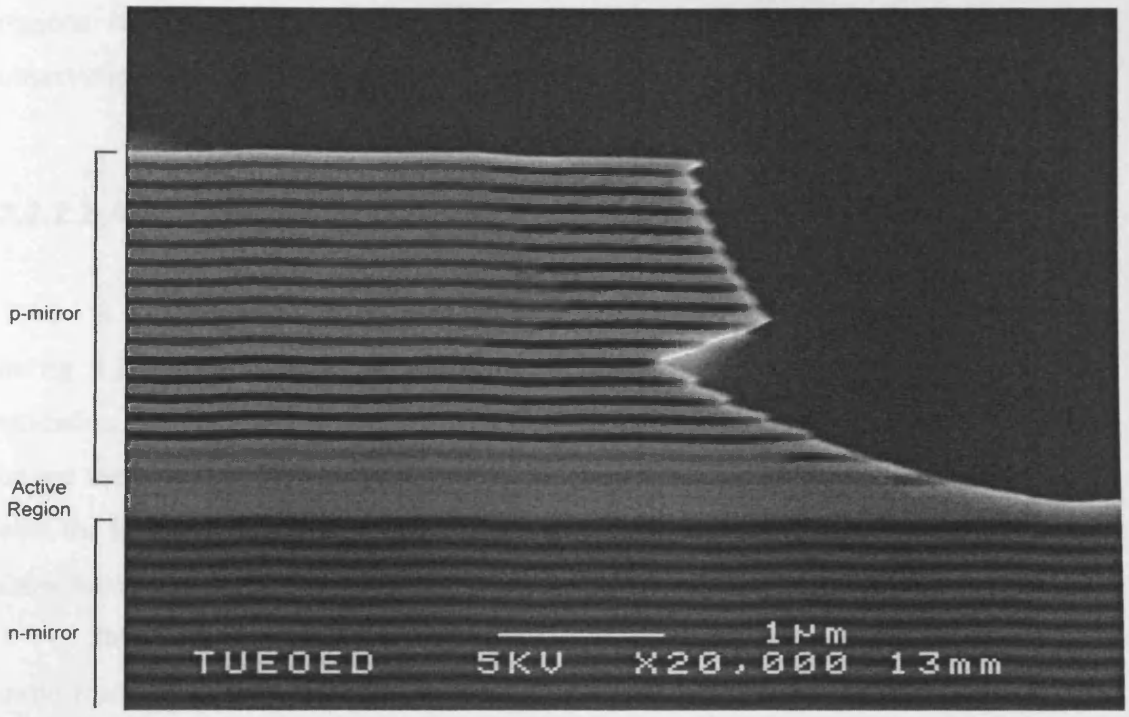

Fig 7.2 *An SEM image of the mesa edge of the VCSEL device QT1500. Layers that have undergone lateral oxidation in the top mirror appear darker.*

The profile of the mesa is clearly discernible in the image, with lateral oxidation occurring from the right hand edge inwards. Individual layers in the two DBRs were distinguished, and the presence of the active region was revealed, although resolution of the quantum wells was not possible. It was observed that the upper portion of the active region had been etched away.

The extent of lateral oxidation for the high mole fraction layers in the upper *p*-mirror is revealed by the darker regions. The appearance of similar dark layers in the *n*-mirror is probably due to the doping difference, since these layers would not have undergone oxidation. The presence of the AlAs layers is revealed by the large extent of lateral oxidation, which carries on beyond the left-hand edge of the image, and measured to be (15.0 ± 0.1) µm from the mesa edge. As expected, the extent of oxidation of the upper ten $Al_{0.91}Ga_{0.09}As$ layers was not as great as that of the five AlAs layers and was measured to be (1.5 ± 0.1) µm. However, the extent of oxidation of the two $Al_{0.91}Ga_{0.09}As$ immediately above the active region is less than that of the upper $Al_{0.91}Ga_{0.09}As$ layers, being (0.5 ± 0.1) µm for the upper layer and (0.3 ± 0.1) µm for the lower layer. It is also noted that the contrast of the unoxidised

regions of these layers is less than the above layers. The significance of this observation was not evident until the AFM investigation was carried out.

7.2.2.2 AFM Analysis of the Sample

A topographic AFM image of the sample after one hour of oxidation is shown in fig 7.3. The orientation is different from the SEM image above, the lateral oxidation in this case occurred from the left-hand side inwards. It was not realistic to image the entire profile of the mesa in this case, because of the difficulties associated with the interaction of the tip with the edge of the sample. However, the image does show hints of the same structure as was seen by the SEM in fig 9.2.

Initial observations noted the expected height contrast between high and low mole fraction DBR layers in the material that had not undergone lateral oxidation, similar to that observed with the VCSEL structures investigated in section 7.2.1. A different topography is observed for regions of layers that have undergone lateral oxidation, caused by the roughness of the porous oxide, revealed by the cleave.

The discussion of the analysis of the AFM is presented below. In this case, the sample has been split up into different layer sections as indicated by the assigned layer numbers indicated in table 7.3.

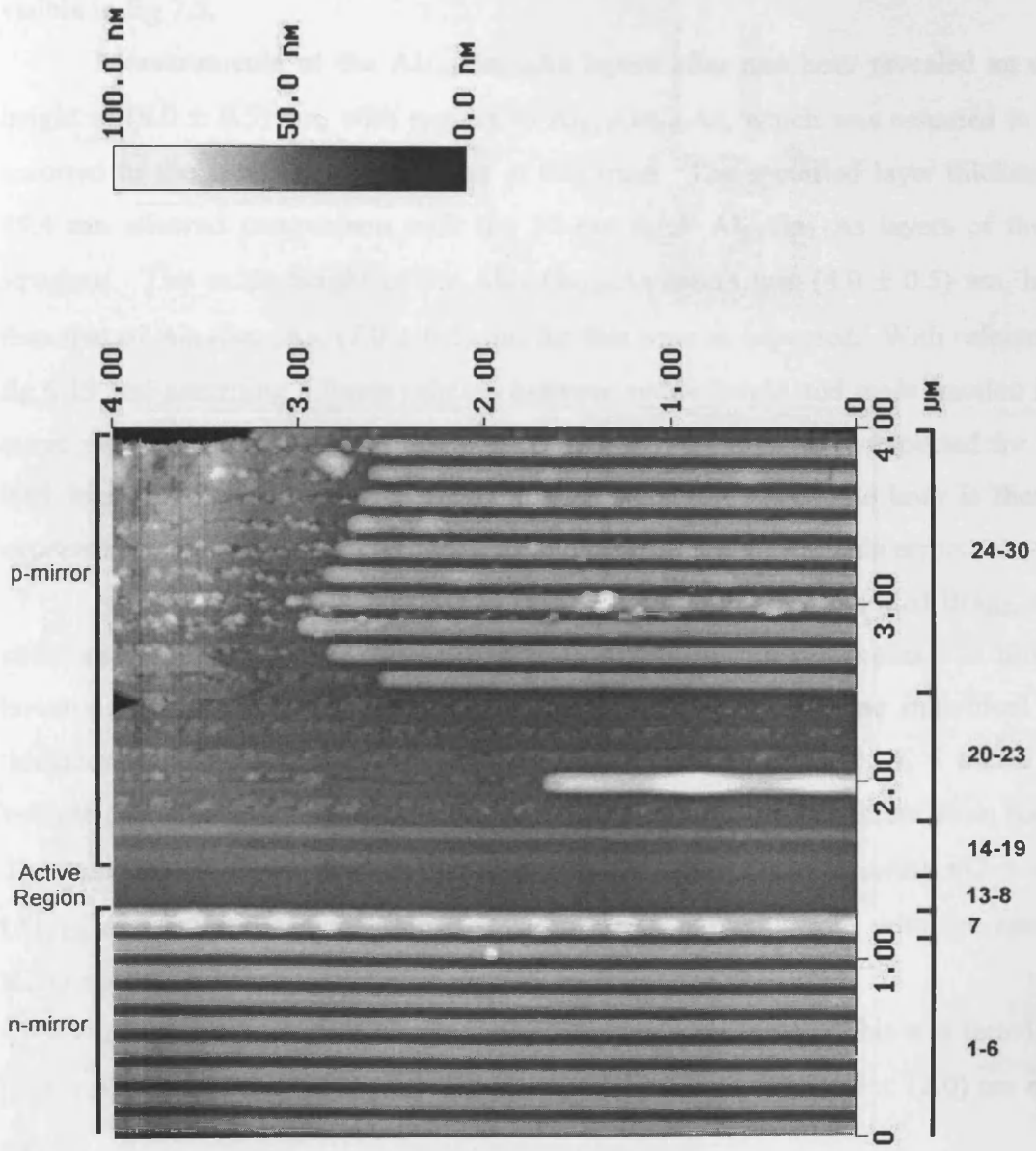

Fig 7.3 A topographic AFM image of laterally oxidised VCSEL mesa device QT1500 after one hour of oxidation. Layer numbers as described in the sample specification are indicated on the right

7.2.2.2.1 Measurements of Layers (1 - 7) x 20 (The Lower Bragg Mirror)

These layers made up the lower *n*-mirror, with layers 3 - 6 being repeated 20 times to form a $Al_{0.91}Ga_{0.09}As/Al_{0.15}Ga_{0.85}As$ superlattice. The upper 8 repeats are visible in fig 7.3.

Measurements of the $Al_{0.91}Ga_{0.09}As$ layers after one hour revealed an oxide height of (8.0 ± 0.5) nm, with respect to $Al_{0.15}Ga_{0.85}As$, which was assumed to have oxidised to the same extent as GaAs at this time. The specified layer thickness of 48.4 nm allowed comparison with the 50 nm thick $Al_{0.9}Ga_{0.1}As$ layers of the test structure. The oxide height of the $Al_{0.91}Ga_{0.09}As$ layers was (8.0 ± 0.5) nm, higher than that of $Al_{0.9}Ga_{0.1}As$, (7.0 ± 0.3) nm for this time as expected. With reference to fig 6.19 and assuming a linear relation between oxide height and mole fraction in the range x = 0.9 - 1.0, an increase in oxide height of 1.5 nm is to be expected for every 0.01 increase in x. The oxide height for $Al_{0.91}Ga_{0.09}As$ after one hour is therefore expected to be 7.5 nm, in agreement with the measured height within errors.

As with the previous sample 1-3280D, the presence of a graded Bragg mirror structure makes layers thickness determination slightly more complex. In this case however, the grade is stepped, and it was possible to determine individual layer thicknesses. The oxidation rate differences between layers 2, 3, 4, 5 and 6 were sufficiently large as to allow their resolution using the profile differentiation method. The measured thicknesses were found to (49 ± 4) nm ($Al_{0.91}Ga_{0.09}As$), (42 ± 4) nm ($Al_{0.15}Ga_{0.85}As$) and (16 ± 4) nm ($Al_{0.48}Ga_{0.52}As$), agreeing well with the specified thicknesses.

In addition, the periodicity of the DBR was determined. This was found to be (124 ± 5) nm, which agrees well with the specified value of (121.4 ± 12.0) nm within errors.

Layer 7, immediately below the active region was also composed of $Al_{0.91}Ga_{0.09}As$. The measured thickness of (65 ± 4) nm agrees well with the specified value, but the measured oxide height of (10.3 ± 0.5) nm is perhaps a little larger than expected for a layer of this thickness. Errors for both these values were inferred from those of the other layers in the DBR.

With reference to the test structure, and the application of a similar argument to that above, it was determined that the layer is in fact composed of $Al_{0.92}Ga_{0.08}As$.

This is further evidenced by the observed presence of inhomogeneities in the oxide growth for this particular layer. It is noted that similar inhomogeneities were not observed for $Al_{0.9}Ga_{0.1}As$ or AlAs layers of similar thickness in the test structure, but have been previously observed in VCSEL structures.

7.2.2.2.2 Measurements of Layers 14 - 19 ($Al_{0.91}Ga_{0.09}As$ Layers above the Active Region)

Examination of the AFM image fig 7.3 indicates these two $Al_{0.91}Ga_{0.09}As$ layers are different from the others. The oxide heights relative to $Al_{0.15}Ga_{0.85}As$ are (1.4 ± 0.2) nm for both the (88 ± 6) nm and (66 ± 6) nm layers. Assuming the effects of oxidation for the surrounding $Al_{0.15}Ga_{0.85}As$ layers is negligible, the mole fraction x, as given by the data in fig 6.19 is x = (0.80 ± 0.02). This does not compare well with the specified value, $x = 0.91$.

7.2.2.2.3 Measurements of Layers 20 - 23 (AlAs Oxidation Layers)

Rapid oxidation of the (49.1 ± 5.0) nm AlAs layers in this region was responsible for producing the desired current aperture. Ignoring the effect of oxidation of the surrounding layers, measurements of regions of the AlAs layers that had not undergone lateral oxidation revealed oxide heights of (16.9 ± 0.4) nm after one hour of oxidation. This oxide height was consistent with oxide heights obtained for the comparably thick 50 nm wide AlAs of the test structure. A thickness of (44 ± 4) nm was measured for the AlAs layers, with the ($Al_{0.15}Ga_{0.85}As$) being (14 ± 4) nm, both agreeing well with the specified values. The periodicity of these layers was measured as (126 ± 5) nm, within the specified value of (122.1 ± 12.0) nm.

Comparisons of the total measured thicknesses for the 5 repeats of layers 20 - 23 for both laterally oxidised and unoxidised regions gave values of (635 ± 5) nm and (630 ± 5) nm respectively. The similarity of these results indicated that no significant distortion of the superlattice has occurred as a result of oxidation of the AlAs.

Lateral oxidation depths of the AlAs layers could not be accurately measured by AFM because of the necessity of displacing the scan region in order to obtain the

depth. However, it should be noted that much higher accuracy may be obtained for measurements of the oxidation depths of the $Al_{0.91}Ga_{0.09}As$ layers, since this can be observed in a single image.

An interesting observation was that not all the AlAs layers laterally oxidised to the same depth, as is clear in fig 7.3. The anomalous layer here oxidised to a depth of only (5.0 ± 0.1) μm, whereas the remaining four layers oxidised to a depth of (15 ± 2) μm, the latter result agreeing with the oxidation depths obtained by SEM. Measurements of the oxidation depths from the opposite edge of the mesa reproduced these results. AFM measurements of these layers had already established confidence in the accuracy of composition and thickness, indicating that the difference in oxidation depth was not caused by errors in these two quantities. This is supported by SEM images of other mesa devices fabricated from the same wafer that show similar oxidation depths for all five layers. The reason for this result is not known at this time.

7.2.2.2.4 Measurements of Layers 24 - 30 (Top of the p-Mirror)

These layers comprised the top 10 periods of the upper *p*-mirror. Measurements of the oxide heights and broadening were in agreement with those obtained for the lower *n*-mirror.

As observed in the SEM image, the extent of lateral oxidation followed the shape of the edge of the mesa. Lateral oxidation depths were measured to be (1.5 ± 0.1) μm, in agreement with the SEM result.

7.2.2.3 Comparison with SEM

It is clear from this comparative study that AFM and SEM compliment each other well as characterisation techniques. They showed much the same features for the sample under investigation. The weakness of SEM compared to AFM is that contrast changes between the layers are not quantifiable in terms of composition. Hence AFM identified two possible deviations from the quoted specification. These findings are qualitatively supported by the SEM image on closer inspection. The determination of layer thickness was also difficult, as the technique was operating

near its resolution limit for semiconductors. However, the profile of the mesa and oxidation depths are more readily obtained with SEM due to the fact there is no solid probe that may be damaged by the interaction with the edge of the wafer. Used in conjunction, these techniques have demonstrated the powerful ability to fully characterise this device.

7.2.3 Investigation of the Bragg Stack of a Light Emitting Diode

7.2.3.1 Background

During this investigation, AFM was used to determine the layer thicknesses in a distributed Bragg reflector where their accuracy was in question. The DBR was one of three light emitting diodes, designed and extensively characterised as part of a postgraduate research project at the University of Manchester, Institute of Science and Technology (UMIST) [1].

The designed periodicity of the DBRs was 1518 Å, each repeat consisting of 580 Å AlAs and 486 Å GaAs layers, with a 226 Å graded region either side of the AlAs layer. Reflectivity measurements, designed to reveal the main cavity resonance, produced a value of 1553 Å for the DBR periodicity. On the other hand, X-ray diffraction measurements produced a value of (1610 ± 27) Å. During the growth of this particular sample, it was suspected that "a malfunction of the MBE equipment may have occurred that could have resulted in incorrect AlAs layer thicknesses being produced".

7.2.3.2 AFM Analysis of the Sample

A topographic AFM image of the oxidised sample cross-section and corresponding line average after an hour of oxidation is presented in fig 7.4. Visible are the top layers of the DBR. It is immediately obvious that four anomalous layers were indeed present. The layers, labelled 1-4, have clearly not oxidised to the same extent as the surrounding layers, implying either a narrower layer, or a lower Al concentration. It should also be noticed that the AlAs layers are round topped rather

than flat or crowned, implying advanced oxidation. Also, the GaAs layers are flat confirming that the AFM has reached the bottom of the oxide trench.

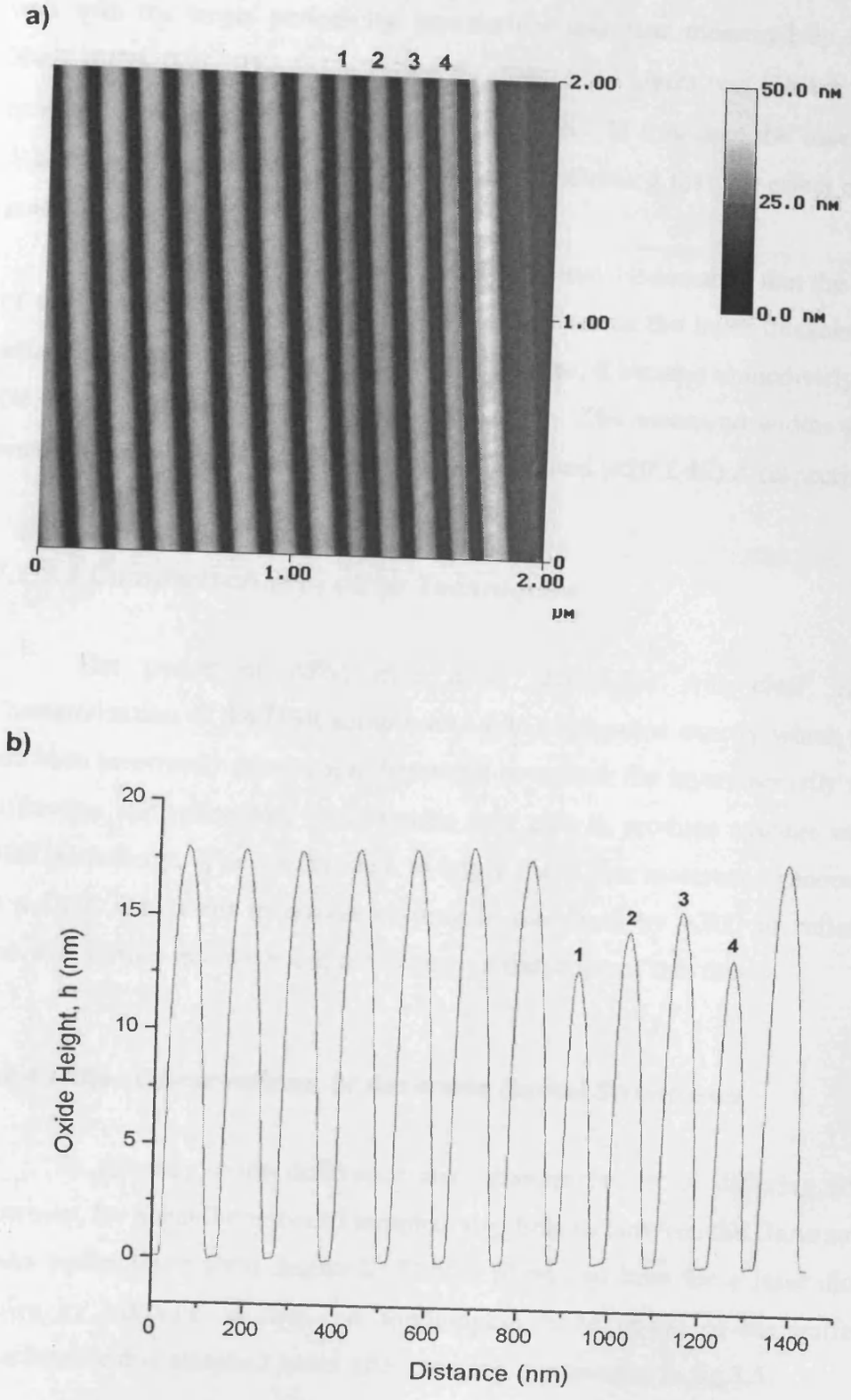

Fig 7.4 Topographic AFM cross-section ***(a)*** *and line average* ***(b)*** *of the DBR structure. The substrate is to the left.*

The periodicity of the remaining layers was measured as outlined for the VCSEL structures earlier, and was found to be (1493 ± 40) Å, a result that agreed well with the target periodicity, but smaller than that measured by reflectivity or X-ray diffraction. The oxide height for these AlAs layers was (16.1 ± 0.5) nm, with broadening an measured with of (600 ± 40) Å. In this case the base width of the differential interface peak was well defined, indicating that the effect of the linearly graded region (226 Å) was minimal.

Following arguments laid out earlier, it may be assumed that the compositions of the AlAs layers are correct, and any errors are in the layer thicknesses. Indeed, when the differential of the oxide layer was taken, it became immediately obvious that the layers in question were indeed too narrow. The measured widths for layers 1-4 were (408 ± 40) Å, (462 ± 40) Å, (500 ± 40) Å and (420 ± 40) Å respectively.

7.2.3.3 Comparison with other Techniques

The power of AFM over other techniques was clear in this case. Characterisation of the DBR sample was able to pinpoint exactly which of the layers had been incorrectly grown, and determine how thick the layers actually were. X-ray diffraction and reflectivity studies were only able to produce average values for the DBR periodicity. The results seem to imply that a few incorrect or anomalous layers in a DBR can result in erroneous periods measured by XRD or reflection. This warrants further investigation, but is beyond the scope of this thesis.

7.2.4 Other Observations of Arsenide Based Structures

In general, oxide difference steps require layers of differing composition. However, for a small number of samples, step heights between the GaAs substrate and GaAs buffer layer were detected. This is illustrated here for a laser diode sample grown by MOVPE at IQE. A topographic AFM image of the buffer/substrate interface for this sample 2 hours after cleaving is presented in fig 7.5.

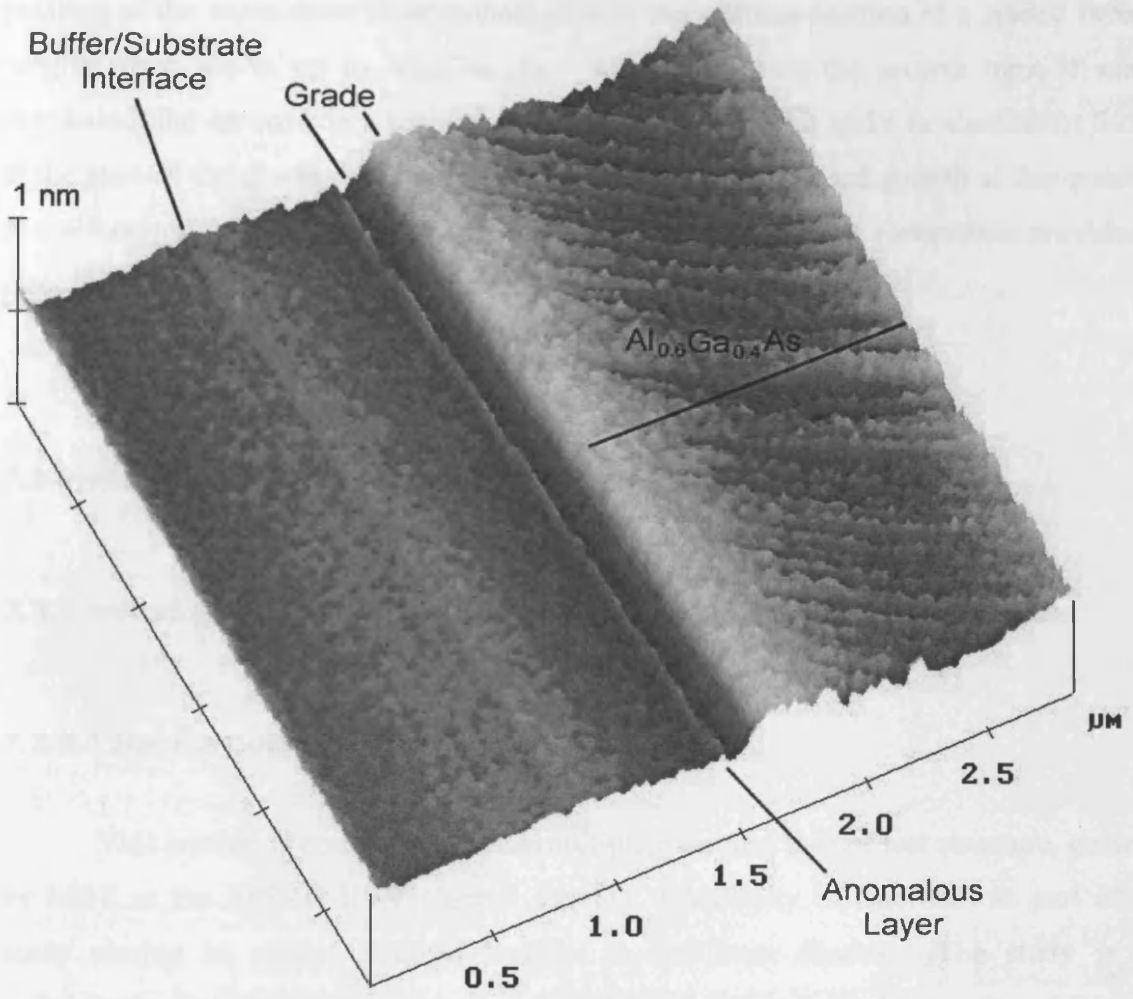

Fig 7.5 *A topographic AFM image of the buffer/substrate interface region for the laser diode after two hours of oxidation.*

A step height of 0.4 Å was observed at the buffer/substrate interface. Both layers were silicon doped, having dopant concentrations of 10^{19} cm^{-3} and 10^{18} cm^{-3} respectively. At this stage, the cause of the observed contrast is not clear. A number of factors could have been responsible. The height difference could be a result of differing oxidation caused by the two different doping levels, or due to the different growth conditions for the substrate and buffer layer resulting in differences in defects and impurities and strain.

Also noteworthy in fig 7.5 is the appearance of the anomalous oxide layer. This is a small (0.2 nm), narrow (50 nm) layer of oxide whose presence was not

expected. The raised layer to its immediate right is composed of $Al_{0.6}Ga_{0.4}As$. The position of the anomalous layer coincided with the starting position of a graded layer ranging from GaAs up to $Al_{0.6}Ga_{0.4}As$. After consulting the growth logs, it was concluded that an error in a mass flow controller produced a spike in aluminium flux at the start of the grade, leading to the observed oxide enhanced growth at this point. A malfunction of the mass flow controller was suspected. This observation provided direct confirmation.

7.3 Investigation of Phosphide Based Devices

7.3.1 Investigation of a Complex Multiple Quantum Barrier Structure

7.3.1.1 Background

This section is concerned with a multiple quantum barrier test structure, grown by MBE at the EPSRC III-V Central Facility, University of Sheffield as part of a study aiming to reduce electron leakage in red laser diodes. The study is a collaboration between the University of Wales Cardiff and the University of Wales Swansea. The structure contains a large number of very thin layers and was produced as a preliminary calibration sample; hence the thicknesses of the layers might be expected to deviate from their expected values.

The sample itself was $(Al_{0.3}Ga)In_{0.49}P/GaAs$ based with three $Ga_{0.51}In_{0.49}P/Al_{0.51}In_{0.49}P$ multiple quantum barrier regions. Two of these regions contained ten 45 Å $Ga_{0.51}In_{0.49}P$ quantum wells, while the third contained ten $Ga_{0.51}In_{0.49}P$ quantum wells in a so-called *chirped* structure, where successive wells increase in thickness ranging from 45 Å to 70 Å. The barrier material was $Al_{0.51}In_{0.49}P$ in each case. Thicker GaAs layers were included between the multiple quantum barriers, providing ohmic contacts that allow the electrical properties of the barriers to be probed.

It was already known that the materials involved had very slow oxidation rates [2], resulting in small oxide height differences. Relative oxidation rates could

therefore be small, reducing the ability of the profile differential method to accurately determine the layer thicknesses. However, it was expected that the extent of any oxide broadening would be small, indicating that reasonably accurate layer thickness determination was possible by base width measurement. The lack of sufficient reference data meant that unfortunately, the determination of composition was not possible.

This is a very interesting sample to measure. The very small differential oxide heights require an excellent sample cleave, but also enhance the chances of measuring the very thin layers that are present. The project allowed for cross-sectional STM measurements of the samples under UHV in Swansea, so a comparison between the two techniques could also be made.

7.3.1.2 AFM Analysis of the Sample

The sample structure is given in table 7.4, along with the AFM measured thicknesses for each of the layers. It was found that the differences in oxidation rates between layers were indeed too small for the profile differential method to be applied. The thicknesses presented are thus the oxide base widths.

In the case of the quantum well structures, uncertainties (outlined below) prevent accurate determination of the individual well and barrier thicknesses. For this reason, the total thickness of the well regions is presented. Errors in layer thickness measurements are assumed to be the same as those derived from the measurement of the standard 2 μm pitch calibration sample, i.e. 2 - 3 %. The extent of oxide broadening is assumed to be negligible due to the small oxide heights.

A topographic AFM image of the cross-section of the complete structure after 2 hours of oxidation is shown in fig 7.6. Fig 7.7 shows detailed images of two of the multiple quantum well regions after the same amount of oxidation.

Table 7.4 *The specification of the multiple quantum barrier test structure with specified layer thicknesses together with those measured by AFM after 2 hours of oxidation. Layer 1 is at the surface.*

Layer	Repeats	Material	Specified Thickness (Å)		Measured Thickness (Å) ± 3 %
1	1	GaAs	5000		4932
2	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	5000		6184
3	1	$Al_{0.51}In_{0.49}P$	283		275
4	1	$Ga_{0.51}In_{0.49}P$	43	Total	1008
5	10	$Al_{0.51}In_{0.49}P$	45	=	
6	10	$Ga_{0.51}In_{0.49}P$	43	923	
7	1	$Al_{0.51}In_{0.49}P$	283		257
8	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	5000		5988
9	1	GaAs	500		626
10	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	5000		5980
11	1	$Al_{0.51}In_{0.49}P$	1489		1487
12	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	5000		6223
13	1	GaAs	1500		1526
14	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	5000		6145
15	1	$Al_{0.51}In_{0.49}P$	283		281
16	1	$Ga_{0.51}In_{0.49}P$	71		1318
17	1	$Al_{0.51}In_{0.49}P$	70		
18	1	$Ga_{0.51}In_{0.49}P$	68		
19	1	$Al_{0.51}In_{0.49}P$	68		
20	1	$Ga_{0.51}In_{0.49}P$	67		
21	1	$Al_{0.51}In_{0.49}P$	65		
22	1	$Ga_{0.51}In_{0.49}P$	63		
23	1	$Al_{0.51}In_{0.49}P$	62		
24	1	$Ga_{0.51}In_{0.49}P$	60	Total	
25	1	$Al_{0.51}In_{0.49}P$	59	=	
26	1	$Ga_{0.51}In_{0.49}P$	57	1208	
27	1	$Al_{0.51}In_{0.49}P$	56		
28	1	$Ga_{0.51}In_{0.49}P$	54		
29	1	$Al_{0.51}In_{0.49}P$	54		
30	1	$Ga_{0.51}In_{0.49}P$	52		
31	1	$Al_{0.51}In_{0.49}P$	51		
32	1	$Ga_{0.51}In_{0.49}P$	49		
33	1	$Al_{0.51}In_{0.49}P$	48		
34	1	$Ga_{0.51}In_{0.49}P$	46		
35	1	$Al_{0.51}In_{0.49}P$	45		
36	1	$Ga_{0.51}In_{0.49}P$	43		
37	1	$Al_{0.51}In_{0.49}P$	283		281
38	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	5000		6184
39	1	GaAs	1000		1056
40	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	5000		6106
41	1	$Al_{0.51}In_{0.49}P$	283		258

42	1	$Ga_{0.51}In_{0.49}P$	43	Total	1005
43	10	$Al_{0.51}In_{0.49}P$	45	=	
44	10	$Ga_{0.51}In_{0.49}P$	43	923	
45	1	$Al_{0.51}In_{0.49}P$	283		264
46	1	$(Al_{0.3}Ga_{0.7})In_{0.49}P$	10000		12400
47	1	GaAs	3325		

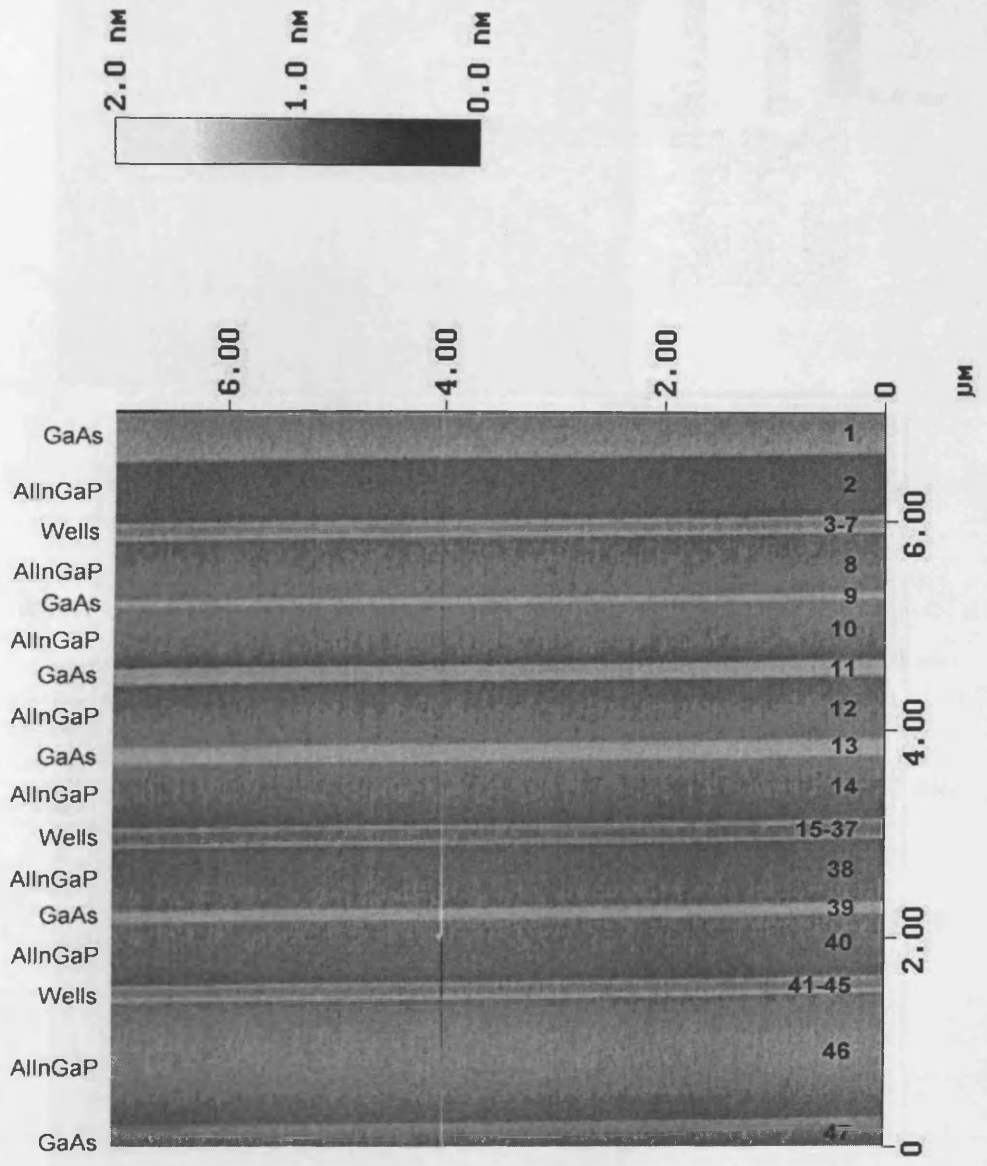

Fig 7.6 A topographic AFM image of the complete multiple quantum barrier structure after 3 hours of oxidation. The surface is towards the top, while the substrate is towards the bottom. Layer numbers, as described in table 7.4 are indicated along the right of the image.

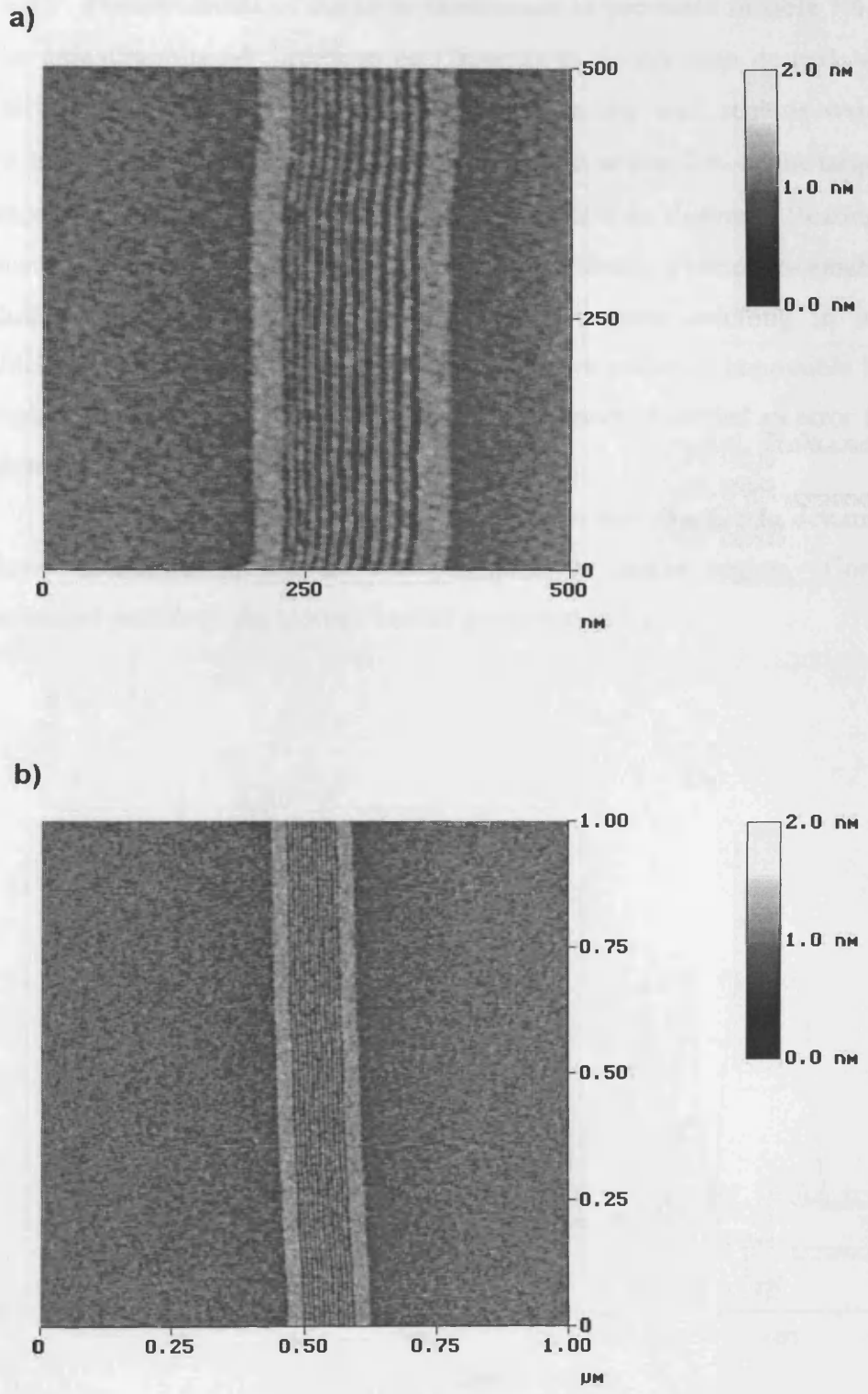

Fig 7.7 Topographic AFM images of the chirped multiple quantum barrier region ***(a)*** *and the upper quantum barrier region* ***(b)****. The effects of drift are evident in image* ***(a)****.*

Measurements of the layer thicknesses as presented in table 7.4 showed all of the $(Al_{0.3}Ga_{0.7})In_{0.49}P$ layers to be (20 ± 2) % thicker than desired, while the total thickness of the $Ga_{0.51}In_{0.49}P/Al_{0.51}In_{0.49}P$ quantum well regions were consistently (9 ± 2) % thicker. All of the GaAs layers were within 2 % of the target thicknesses, except layer 9, which was (25 ± 2) % thicker than desired. Bearing in mind the accuracies quoted from previous studies in this thesis, it seems reasonable to conclude that there were errors during the growth process resulting in incorrect layer thicknesses. However, the distribution of errors makes it impossible to speculate of their source. It would have been satisfying to have identified an error for a particular element (Ga, In or Al).

It was mentioned previously that it was not possible to determine individual layer thicknesses within the multiple quantum barrier region. Consider the line averaged profile of the chirped barrier presented in fig 7.8.

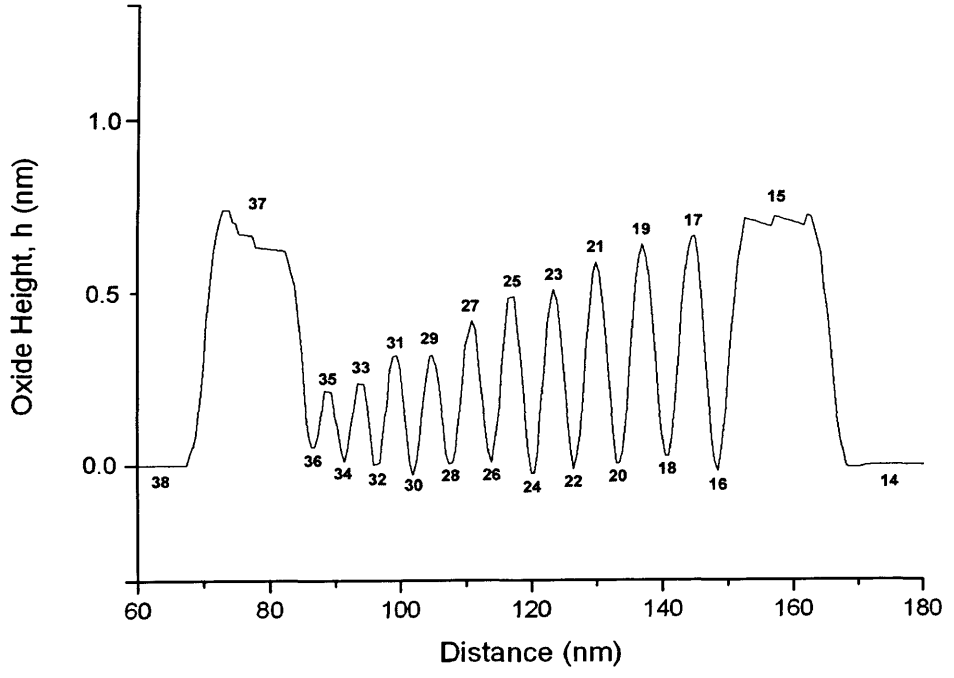

Fig 7.8 A line averaged profile of the chirped multiple quantum well structure. Layer numbers are indicated. The wells and barriers become wider on going from left to right

These effects arise from spreading of the oxide, tip shape convolution effects and layer width dependent oxide heights. The small height of layer 35 compared to the other layers is thought not to be due to tip convolution, and is assumed to be due to layer width dependence. In this case it is also noted that the variation in height of the layers is consistent with a tip radius of ~ 20 nm. Further evidence for the effects of tip convolution is the fact that the base level of trench bottoms of the narrower layers is higher than the remainder of the layers. However, the effects of oxide broadening cannot be ruled out in this case.

7.3.1.4 Oxide Heights for Phosphide Materials

This section outlines typical oxide heights obtained from the phosphide based materials within this structure as a guide for future study. Other work has shown that $Ga_{0.51}In_{0.49}P$ oxidises more slowly than GaAs [2], but it has not been able to successfully measure oxide height differences for this sample due to the effects discussed earlier. A tip convolved oxide height difference of just 0.1 Å was measured between layers 35 and 36 after one hour of oxidation. This height is the smallest possible *z* deflection for the AFM scanner, and is at the limit of resolution for the optical lever method [3].

However oxide heights of (-0.7 ± 0.5) Å and (-0.8 ± 0.5) Å have been measured for $Al_{0.51}In_{0.49}P$ material with respect to GaAs after one hour and one day respectively. In the case of $(Al_{0.3}Ga)In_{0.49}P$ with respect to GaAs, heights of (-1.4 ± 0.5) Å and (-2.3 ± 0.5) Å have been measured for the same times.

The oxidation rates of these materials is clearly much slower than that of GaAs, with $(Al_{0.3}Ga)In_{0.49}P$ oxidising the slowest. Not surprisingly, it contains the least amount of Al.

7.3.1.3 Comparison with STM

Scanning tunnelling microscopy (STM) under UHV was conducted on the multiple quantum barrier regions of the sample by the Semiconductor Interface Laboratory, University of Wales Swansea. The STM images obtained of these regions are shown in fig 6.7.

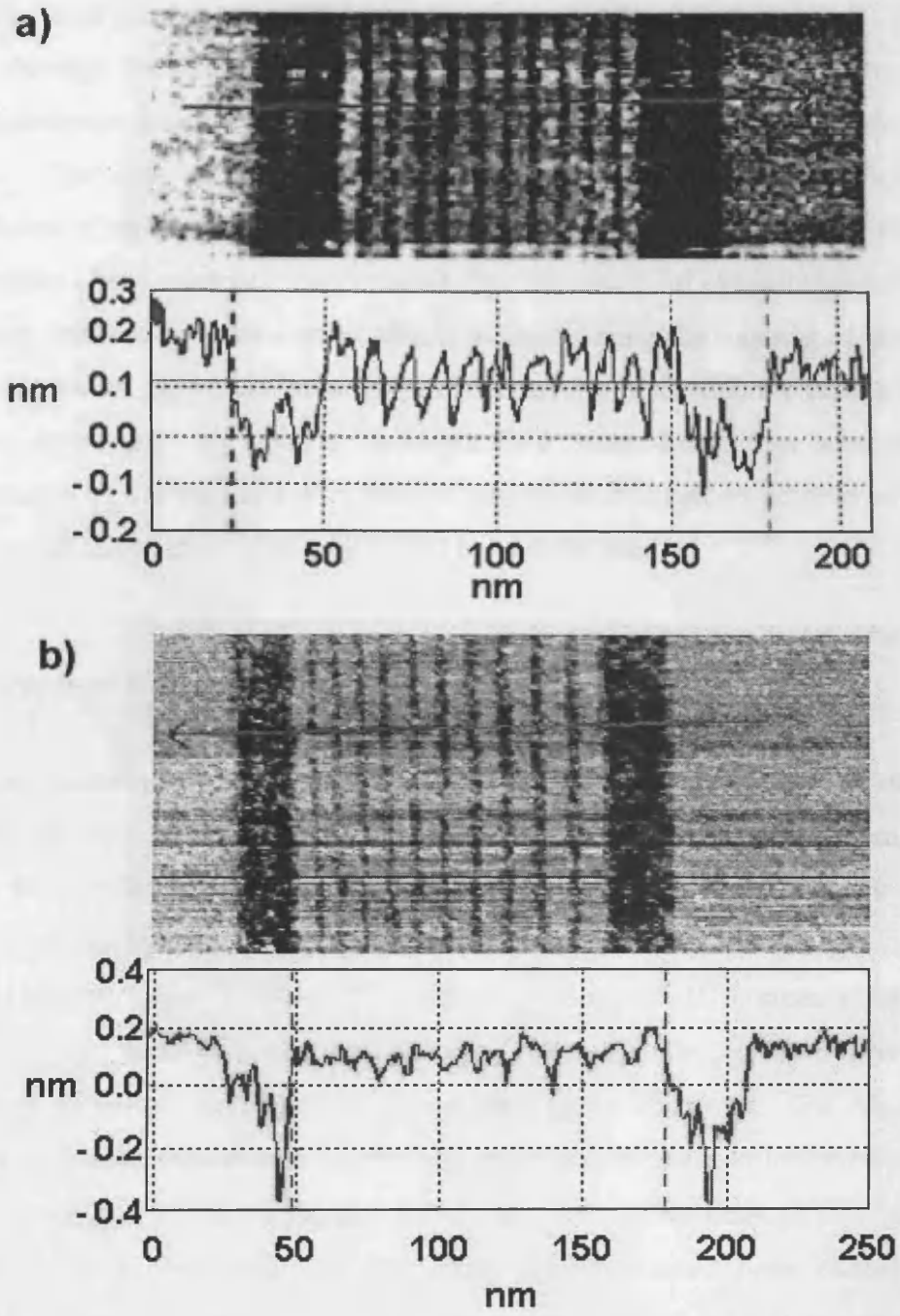

Fig 7.9 *STM images of a multiple quantum barrier region* ***(a)*** *and the chirped multiple quantum barrier region* ***(b)****. The thick dark layers are the 283 Å AlInP layers. Dark layers within the well regions are also AlInP, while the lighter layers are GaInP. The arrows indicate the length along which the cross-sections (below) are taken. Plots below show line sections as indicated on the STM images. Contrast here is produced in differences between the bandgaps of the materials.*

The data obtained by STM is at best comparable to that of AFM. The line sections through the STM data reveal ambiguities in measurement that make the accurate determination of layer thickness for the multiple quantum barrier layers difficult. The lack of well-defined interfaces in this case was explained by interdiffusion of atoms at the interface. In the case of AFM, accurate layer thickness determination was clearly compounded by the lack of knowledge of oxide broadening, and the tip convolution effects evident during the imaging of the barrier regions. However, AFM has been able to successfully determine the existence of all the layers, despite the very small oxide height differences. In addition, layer thickness measurements of the thicker layers would have been difficult to achieve with STM because of the large image sizes that would have to be used.

7.3.2 Investigation of a Laser Diode

One issue briefly addressed previously in this study is the effect of strain and strain relaxation in epitaxial layers. A small number of samples have been studied that may show related effects, and one of these is presented here. Fig 7.10 shows a laser diode grown by MOVPE at IQE.

At the top was a 2.75 μm thick $Al_{0.7}Ga_{0.3}As$ layer. Its bottom interface was with a 250 nm thick $(Al_{0.7}Ga_{0.3})In_{0.5}P$ layer. The remaining lower layers were $(Al_xGa)In_yP$ of various composition, grown on a GaAs substrate. The $Al_{0.7}Ga_{0.3}As$ layer was sufficiently thick as to be partially relaxed in the case of lattice mismatched growth. Striations of amplitude (0.7 ± 0.2) nm are visible through this layer that terminate at its bottom interface and could therefore have been related to the associated dislocation network. Striations are present through a DBR on one side of a VCSEL in fig 7.1, which given the large $Al_xGa_{1-x}As$ content, may also be strain relaxation related.

Similar striations have been observed previously [4] during cross-sectional studies of AlSb based device structures carried out by AFM in air. In this case, TEM had indicated the presence of the striations before the sample was exposed to air. The striations were attributed to a lattice constant mismatch between AlSb and GaSb (0.656 %). The same work did not rule out the effect of dopants within the structure, however.

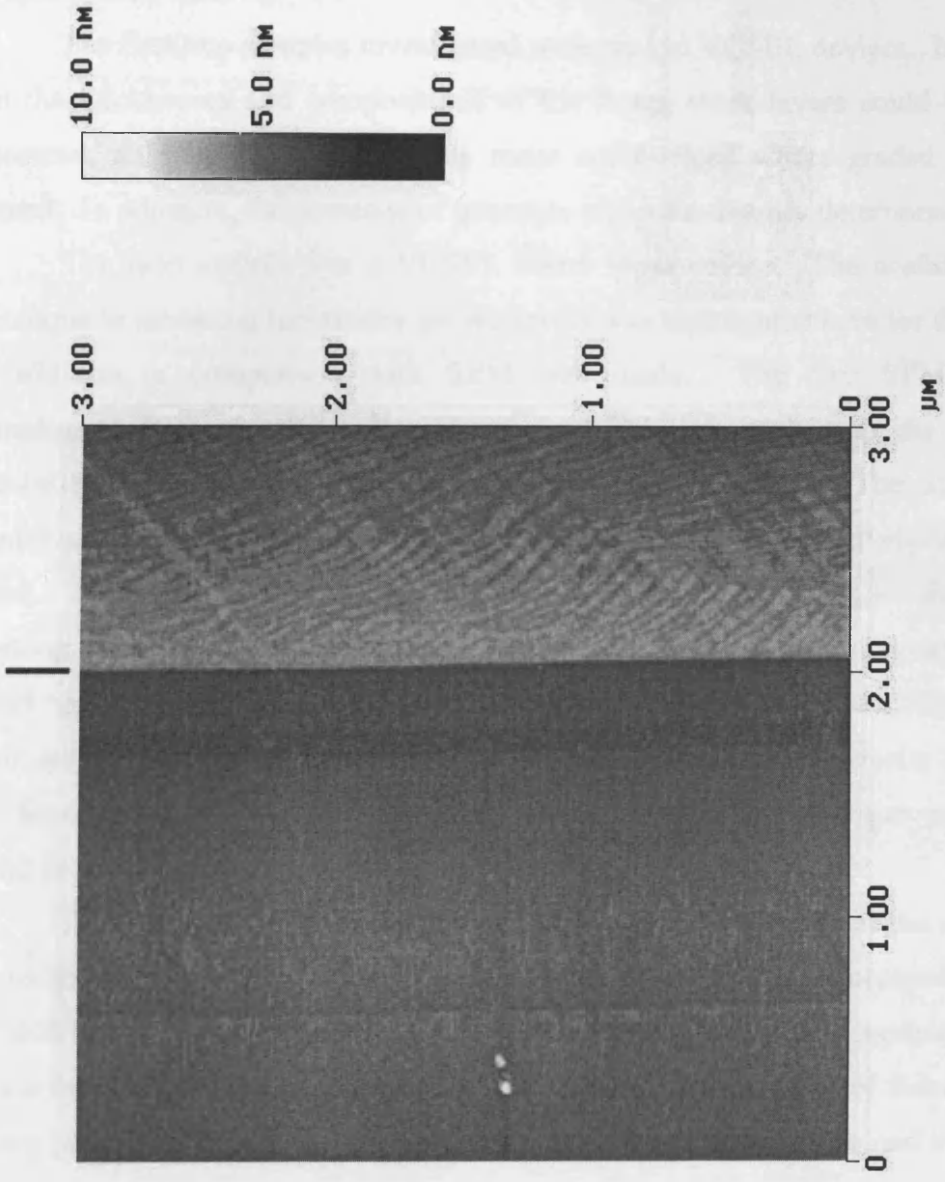

Fig 7.10 *A topographic AFM image illustrating striations present on a $Al_{0.7}Ga_{0.3}As$ layer grown on $(Al_{0.7}Ga_{0.3})In_{0.5}P$. The interface between these layers is indicated.*

7.4 Summary

Throughout this chapter, several contrasting studies have been presented that will hopefully give the reader an appreciation of how powerful this technique can be. It has been used both on its own, and in conjunction with other SPM techniques in a complimentary fashion.

The first two samples investigated were typical VCSEL devices. It was shown that the thicknesses and compositions of the Bragg stack layers could be routinely measured, although this was slightly more complicated where graded layers were present. In addition, the presence of quantum wells was readily determined.

The next sample was a VCSEL based Mesa device. The usefulness of this technique in revealing incorrectly grown layers was highlighted here for the first time. In addition, a comparison with SEM was made. The two SPM techniques complimented each other well. However, SEM was pushed to the limit of its capability, and showed some weaknesses compared to AFM. The deviations the sample showed from the sample specification were only qualitatively supported by SEM. The determination of layer thickness was also difficult, as the SEM was working at the limit of its resolution. Finally, the observed contrasts between layers could not be quantified in terms of composition, which is possible using AFM. The main advantage SEM had over AFM in this case was that it was easier to determine the lateral oxidation depth of the current aperture, due to the larger scan size that could be achieved.

The light emitting diode sample showed a further example of the usefulness of the technique to specifically pinpoint incorrectly grown layers. A comparison with X-ray diffraction and reflectivity studies of the Bragg stack of the sample was made. While these two techniques were able to determine the presence of these incorrectly grown layers, AFM was able to identify exactly which layers contained the error, and specifically that it was an error in layer width.

The study of the final AlGaAs based sample was included as a further example of how useful the technique is for highlighting sample defects. A height difference between substrate and buffer layer was discovered, along with an anomalous layer whose presence was not suspected previously. Both of these were immediately observable from the AFM images obtained.

The first of the phosphide samples discussed pushed the technique to its limit in terms of measuring height and thickness. This multiple quantum barrier structure contained very narrow layers of very slowly oxidising material. AFM was able to detect all the layers, with measurable oxide height differences between each. A direct comparison with STM was made and the results were found to be comparable. While atomic resolution is possible with STM, the lack of well defined layer interfaces in this case meant that AFM was able to more accurately determine the individual layer thicknesses.

Finally, the last phosphide based sample provided direct observation of strain induced striations within a device. The ability to detect these was unhindered by the oxide growth on the surface.

7.5 References

[1] M de la Fargue. PhD thesis. University of Manchester, Institute of Science and Technology.

[2] C W Wilmsen (Ed.). *Physics and Chemistry of III-V Compound Semiconductor Interfaces*. Plenum Press (1985).

[3] *MultiMode™ SPM Instruction Manual*. Digital Instruments.

[4] D Convey & R Legge. Microscopy and Analysis, 19 (November 2000).

Chapter 8

Discussion and Suggestions for Future Work

8.1 Discussion

The main objective of this work was to assess the capability of the AFM to perform accurate metrology of III-V semiconductor device structures, and determine its potential use as a characterisation tool in an industrial setting. This has been achieved to a great extent.

In order to be able to perform cross-sectional AFM on the devices, it was necessary to be able to produce cleaved surfaces of a sufficiently high quality. It was discovered that cleave damage steps of as little as 3.4 Å could render the measurements of oxide step heights impossible. To this end, a relatively simple and effective technique for manually cleaving the wafers was developed, capable of producing cleaves with regions of the order of microns free of cleave damage. A series of experiments were conducted which showed that provided the wafers were thinned down to ~ 150 μm, at least 90 % of cleaves were suitable for AFM analysis, including those on (100) wafers mismatched by as much as 10°.

Oxide step height was measured with respect to GaAs for the $Al_xGa_{1-x}As$ compounds in the test structure and plotted as a function of time. These confirmed the dependence of oxidation rate on layer thickness and composition, with enhanced oxidation occurring for thicker layers and higher aluminium concentrations. Observations indicated that for the time scale of the measurements, field-aided oxidation was the dominant form of oxidation for the compounds $x < 0.8$, while for the materials $x > 0.8$, diffusion limited oxidation dominated.

The evolution of the oxide profiles was also studied carefully. A three stage temporal evolution of the profiles was observed. Initial oxide films have flat tops. For $x = 0.3$ - 0.8 after a period of time, this evolves to a so-called crown topped profile. The crown topped profiles were proposed to be a result of some form of interface enhanced oxidation that eventually merged together producing the rounded topped profile. The observation of crown topped profiles has not been noted in any

previous research. Further oxidation resulted in the emergence of a profile with a rounded top. Fast evolution through these profiles was favoured by high mole fractions of aluminium

Layer widths could be determined easily by differentiating the oxide profiles obtained by topographic measurements. The quantification of oxide step heights for each composition and layer thickness enabled the determination of layer composition for real device structures. Using the test structure as a reference, it was possible to determine layer compositions to within $x \pm 0.02$ and layer widths to within 3 % at best (i.e. the confidence in the growth of the test structure).

From the data collected from the test structure, it was determined that the largest changes in oxidation rate occur in the range $x = 0.8$ - 1.0. Unfortunately, the compositions in the test structure did not allow for a full determination of oxide heights and broadening within this range. It would therefore have been useful to have additional test structures from which data for these compositions could be collected.

Studies of real devices revealed the true potential of the technique. Measurements of VCSELs showed the ease at which thin quantum wells could be detected. The determination of layer thickness for the individual DBR layers was complicated when graded layers were present within the superlattice. However, the DBR periodicity could be determined, even if a considerable amount of oxidation had taken place.

This technique also allowed the determination of oxide depths for a laterally steam oxidised VCSEL mesa structure. A comparison of SEM data from an identical device was made. Although superficially similar, the ability to quantify the observed contrast between the layers in terms of composition makes AFM a much more powerful diagnostic tool. In addition, the thicknesses of individual layers was difficult for SEM due to its limited resolution.

Metrology of phosphide based structures has also been demonstrated, where material oxidation rates were known to be very small. Unfortunately, the lack of a test structure similar to the $Al_xGa_{1-x}As/GaAs$ based test structure, meant that layer thickness or composition determination is not yet possible for these materials.

Measurements of a phosphide based multiple quantum barrier test structure stretched the technique in terms of height differences and layer widths. For this sample, the presence of a quantum well with a measured width of only 42 Å was

detectable from an oxide height of just 0.1 Å. However, it is likely that the effects of tip convolution played a role in the image obtained from this sample.

The effect of tip convolution on the ability to accurately determine the widths and depths of trench-like features, such as those formed by two adjacent layers with oxide films, has also been considered. If the separation of the oxide layers is comparable to the radius of the tip (~ 20 nm), then the tip would be unlikely to penetrate the full trench depth. For the case of thinly oxidised materials, this had a clear effect on the ability to accurately determine oxide step heights between layers, layer thickness and composition.

Measurements of the $Al_xGa_{1-x}As/GaAs$ test structure revealed a retarding effect on the oxide growth during the continuous scanning of a particular region. This was greatest for the AlAs layers, with a retardation in oxide height of 5 % being noted. The extent of retardation decreased with decreasing aluminium concentration for the layers. This observation is consistent with the tip wearing down the (porous) oxide, the porosity of the oxide being greatest for the AlAs material. Further evidence for the increased porosity of AlAs comes from the observation that the structure of the oxide associated these layers began to disintegrate after a few days of oxidation, which was not observed to occur for the other layers of the test structure.

Topographic differences between layers were used to detect their presence. It was found that phase imaging could be used for this purpose when there are no oxide step heights present. This technique confirmed the presence of the $Al_{0.1}Ga_{0.9}As$ layers at a time soon after the cleave, where no discernible oxide layer height was measured. The phase contrast produced here was thought to be due to differences in 'stickiness' between the $Al_{0.1}Ga_{0.9}As$ oxide and surrounding GaAs oxide. Such a result has not been observed before.

In conclusion, it has been demonstrated that AFM can be used very powerfully as a metrology tool for III-V semiconductor devices, despite some minor drawbacks. The ease of sample preparation, combined with relatively quick sample turn around time indicates that it is possible that the techniques outlined in this thesis may be employed in an industrial environment. Indeed, an AFM has been recently acquired by International Quantum Epitaxy Ltd. The justification for the purchase was based largely on the findings of this work.

8.2 Suggestions for Future Work

The most immediate task for any future work is to build up a more complete picture of the oxidation behaviour of $Al_xGa_{1-x}As$ within the range $x = 0.8 - 1.0$. A lack of knowledge of the oxidation of these compounds has resulted in uncertainties in the determination of thickness and composition for layers containing them. The design of new test structures should therefore be considered. Thorough analysis of these structures would lead to information similar to that obtained from the test structure designed for the purpose of this research.

More work on different material systems such as phosphide based materials would be of great benefit. It was only possible to obtain preliminary oxide height measurements for these materials with the limited amount of devices available. The production of a phosphide based test structure similar to that used to characterise the $Al_xGa_{1-x}As$ materials would clearly be beneficial, but the larger lattice mismatch present in these materials would limit the range of compositions that could be investigated.

The oxide enhanced layer contrast mechanism utilised here is not the only way contrast may be achieved between the layers of superlattice structures. In principle, it is possible to use a selective etch to remove material from the cross-sectional surface. The resulting etch depths could be used to determine layer composition, while layer thickness measurements may be more accurate than those obtained here. However, preliminary experiments of this type were carried out at the start of this project without significant success.

It is also possible to probe the electrical properties of the devices by using AFM based techniques such as electric force and scanning capacitance microscopy. These two techniques are becoming increasingly used to determine doping concentration of layers and the position of *pn* junctions, for example. Here, electrical data may be collected simultaneously with the topographic data, allowing comparison between the two. This technique may be compatible with the measurement of oxide enhanced device cross-sections, but it not certain to what extent the oxidation of the surface will complicate the electrical measurement of the sample.